

INSUBSTANTIAL SURVIVAL

PERSONS AND LIVES

by

ERIK DYRHAUG



THESIS

presented for the degree of

MASTER OF PHILOSOPHY

Department of Philosophy, Classics, History of Art and Ideas

Faculty of Humanities

University of Oslo

Fall 2011

Abstract

There are many competing accounts of what a person's identity over time consists in. This thesis addresses this topic, but not this question in particular. Like Parfit, I argue that when it comes to our survival, identity cannot be what matters. I try to make the case that creating reductionist accounts of our persistence is a sidetrack to a more interesting problem. When considering ourselves and others in a situation, we automatically apply the idea of a subject. It is also from this perspective that the question of survival seems most important. Under this conception of ourselves, however, we should not think of ourselves as persisting (unless we are prepared to think of ourselves as being souls). Upon realizing this, we become free to investigate the nature of our survival, without the logic of identity restricting us. From this understanding I try to create an account of personal survival, uniting the idea of a subject with that of a life. I also look at how some of the concepts and practices usually thought to depend on personal identity over time (e.g. desert) can be accommodated within this picture.

Preface & Acknowledgements

I have wondered about the questions in this thesis for several years now, and – as is typical of philosophical questions – I have not reached any final answers yet. I do feel, however, that I have come quite a long way. I have changed my views on these questions quite a few times (which is always exciting), but I have also felt a gradual settling towards the view that I now hold. At this point I see the core of my view as being quite clear, but at the same time quite hard to express. I feel I might be missing a larger and clearer defined framework to express myself in. At the same time, I fear that a larger framework could easily complicate things; make it harder to discuss these questions with whomever I talk to, and make it easier to forget what is interesting about them in the first place.

For helping me along in writing this thesis, I am most of all grateful to Anders Strand, supervisor extraordinaire. All the hours I have sat in his office, hashing out the various aspects of this thesis, all his helpful comments, and all his encouragement; I could not have written this without him.

I am grateful also to Jonas Nordkvelle, not only for reading a draft of this thesis and for giving me helpful comments, but for the countless hours we have spent discussing random topics with no clear goal in sight.

I am grateful to friends and family, for their patience and support.

I am grateful to all the lunatics in Room 109. They have made the last period of writing, not only bearable, but frolicsome and gay. I miss the routine already.

Finally, I am very grateful to Henriette, for her love, patience, and support.

Table of Contents

1	Introduction	1
1.1	What and Why	1
1.2	The Goal	2
1.3	Background Assumptions.....	3
2	Getting at the Problem.....	7
2.1	The Simple Intuitive View	7
2.2	The Simple View Unravelling	8
2.3	The Many Ways of Framing the Discussion	11
3	My View.....	15
3.1	My Life and Me.....	15
3.2	Interpreting the Thought Experiments.....	27
3.3	A Grand Finale	37
4	Competing Accounts	41
4.1	Parfit	41
4.2	Four-Dimensionalism	44
4.3	Animalism	48
4.4	The Psychological Approach.....	52
4.5	The Soul Theory and Other Non-Reductive Views.....	54
5	Consequences of My View	56
5.1	Conceptual Revision.....	56
5.2	Practical Adaptation	58
6	Objections.....	65
6.1	My Premises Are False (or Implausible).....	65
6.2	My Conclusions Are Unacceptable	65
6.3	A Deeper Truth to Be Found	66
6.4	Are Persons ₁ Temporally Extended or Not?.....	67
6.5	There Are Better Accounts	68
7	Conclusion.....	69
7.1	What I Have Tried to Accomplish.....	69
7.2	Further Impact	70
	<i>Appendix : Enumerated Thought Experiments.....</i>	<i>73</i>

A1 Surviving Accidents	73
A2 Suspended Life	74
A3 Replacing Parts.....	75
A4 Fission	77
A5 Beyond.....	79
Bibliography.....	85

1 INTRODUCTION

1.1 What and Why

This thesis concerns how we survive. Not in the sense of what we eat or how we avoid falling down staircases, but in the sense of what it means for us to have a future, and what it is about the future of one of us that makes it that particular person's future. Another way of putting the question is "what relation must you have to some future person, in order for that person's present to be your future?"¹ Is it enough that this person is similar to you in certain relevant ways? That he has the same personality as you, say, or the same good looks? A lot of different factors go into making you the person you are. Your temper, your interests and preferences, your way of thinking about things – and of talking, walking, smiling – your habits and your quirks. All of these features, and all the ones like them, we can think of as constituting your character. But there is more to you than your character. There is also your memories and your projects, your hopes and dreams, your plans for the future. Then there are all the external factors, like your relations to friends and family, your job, and the place you live. But what if all these things about you were to change? Would you be gone then, or is there more to you still? If you think you have something like a soul, you might think that even though all your features, properties and relations were to change, you would still be there, just with different features, properties and relations. And if you do not think you have a soul, you might feel this way all the same – but if you do not have a soul, how could that be?

In our lives we experience changes in ourselves and in people around us. Some of these changes can be so large, so fundamental or so sudden that we might find ourselves pondering whether the old person is still there. If somebody changes enough, we might say things about him like "he has become a completely different person", but most of us do not take such claims to be literally true. What I mean by this is that most of us still judge the fate of the old person by the life of the new, in a way that we would not have, had we taken seriously the idea that this was in fact not the same person. So a person can completely change character, he can completely lose his old memories, and in a sense become a completely different

¹ The more familiar way of posing this question is "What would it take for some future person to *be* you?" But I will argue that for a central way in which we think about persons, a person is essentially *of* the moment, and under this conception a person is never identical to any person at any other time. This conception of "person" will be further probed (especially in Chapter 3.1).

person, but still we would think that he in some way remained. Becoming a different person would in this way be seen as a way of surviving, not of dying.

We are conscious beings that act and experience, we are sets of molecules, and we are biological organisms, we are characters and personalities, and we are persons. But when we hope to survive, what is it we are hoping for? And when we think about our own future, what is it about ourselves that we imagine reaches it?

1.2 The Goal

What I aim to accomplish with this thesis is to hone in on a picture of how we can think about ourselves and our lives that is more in accord with our current scientific perspective. Ideally this picture will be a consistent system describing the different conceptions we use in thinking about ourselves, as well as how these conceptions are related.

To a firm believer of eternal souls, the atheist picture might seem quite surd. And even though I am a non-believer myself, I actually share this sentiment. The story for the atheist goes something like this: We come into existence when we are born, we live for a small number of years, but then, as we die, we go out of existence again. For all the time the universe has existed and for all the time the universe will exist hence, of all those years and all those places, you only get to experience this tiny window right here and now. You are an experiencing subject with a window of experience that starts at some time around birth and ends when you die. From a perspective like this, people think things like: “I’m lucky to have been born here”, or “I’m lucky to live in these times of great scientific discoveries”, or even “He should have lived in the time of the Vikings – he would have been a highly regarded swordsman”. Perhaps we do not really mean these things literally. Still, we seem to think there is something contingent about where our window of experience ended up. Additionally, many philosophers, and mortals alike, have expressed their future non-existence, the absence of any possibly future experience, to be precisely what makes the thought of dying so dreadful.²

² As for philosophers, Nagel, for instance, says this about the evil of death: “[...] he finds himself the subject of a *life*, with an indeterminate, and not essentially limited future. Viewed in this way, death, no matter how inevitable, is an abrupt cancelation of indefinitely extensible possible goods.” (Nagel 1991, 10)

If I am right to suspect there is something wrong with our conception of ourselves and our lives, a better picture might allow us to think more clearly about matters of life and death, about the value of our own futures relative to that of others, and about how we should frame our hopes for the future.

My approach in this thesis is primarily to look at inconsistencies in how we intuitively think about ourselves and the world, and inconsistencies between those conceptions and our current scientific understanding of the world. The goal is to eliminate these inconsistencies by revising our understanding, without giving up more than we have to. The starting point is the barest intuitive grasp we have of ourselves as well as our current scientific understanding of the world. I take this most basic understanding of ourselves to be what is captured by folk psychology.

1.3 Background Assumptions

Whenever you try to say anything you have to take a lot of other things as given. If you are not impressed by science, for instance, you will not be impressed with arguments based on scientific findings. The whole argumentation of this thesis is framed in a physicalistic world view, and what is at stake depends on what kind of a question I think I am posing. These two assumptions will get some attention next.

1.3.1 Physicalism

When I say I assume a physicalistic world view, I do not mean to say that the only real or useful explanations and accounts are physical ones. What I want to express is the assumption that if a situation is held physically fixed, all that can vary is the description of it. I assume that there are no non-physical aspects in the world that interact with the physical ones. One way of understanding this is that I simply *call* everything physical, or else that I call everything physical that can interact with what I know to be. To this I would only add that there are both physical and non-physical descriptions, notions and explanations. I do not take everything to be physical things. Take the number one as an example. It is a number, and so it is *something*, but I do not think that the number one *is a physical thing*. The intention I have of completing this thesis is also *something*, but it is not a physical notion. That is, its not a notion that implies the physicality of what it denotes. I do however assume, and I think with

good reason, that this intention has a physical basis in my head. While there might not correspond a neat physical object to the notion of an intention, the intention is not something operating above and beyond the physical world.³

Physicalism is sometimes defined in terms of supervenience. This is then expressed as the claim that facts of other domains supervene on the physical facts, either locally or globally.⁴ What I am specifically assuming for this thesis is not global supervenience, but rather only that persons are such that if we hold a complete physical description of a person fixed, nothing about that person can be changed without changing his environment. Given that I am myopic, and that I am now smiling and thinking about the cup of coffee I am about to have – we could not hold a complete physical description of me fixed but change the fact that I am myopic, that I am smiling, or even that I am thinking about coffee.⁵ If I had been thinking about something else, something would have been physically different as well – something else would have been going on in my head. Now this, once again, is not to say that we can deduce – even in principle – from a complete physical description of me what I am thinking about. It might be possible, but it *being* possible is a further claim, and it is not one I am making.

1.3.2 Metaphysics and Truth

The topic of personal identity within philosophy is couched within the branch of philosophy called metaphysics. But what is metaphysics? Different philosophers have different ideas about this. Some think metaphysics is the study of how the world is at its most fundamental level. Some (me included) think it is rather the study of how to best make sense of the world as it appears to us. Some draw a line *within* metaphysics, saying that while some questions just concern how we think about the world, others are deep questions about how the world

³ My view is close to something like Papineau's "Conceptual Dualism" (Papineau 2004, 47-72), or Perry's "Antecedent Physicalism" (Perry 2003, 26-92)

⁴ For a rundown of different types of supervenience see (McLaughlin 2010)

⁵ Careful readers might have noticed the "without changing his environment" clause in the previous sentence. It is there to pre-emptively ward off a potential objection to the claim that I cannot even be thinking *about* something else. The objection I have in mind is the one from externalism about mental content. The idea is that if my environment had been relevantly different I would have been thinking about something else even though I would have internally been in the same state. See (Burge 1986) and the original twin-earth thought experiment of (Putnam 1973)

itself is.⁶ What view one has of metaphysics influences the way one attacks metaphysical problems, and indeed what problems one thinks there are to attack in the first place.

I take the proper role of metaphysics to be that of investigating our basic concepts upon which all our further thoughts about the world rest. Typical examples of metaphysical topics include the concepts of objects, of properties, of space, time, events, and so on. A brilliant example of successful conceptual revision (from physics) is Einstein's realisation that space and time are more properly thought of as a single entity: *space-time*. This revision has taken hold within physics, which is where it was aimed at. Outside physics, however, the common-sense notions of space and time prevail – *absolute* space and time even. When it comes to our concepts of persons, agency, or morality, these are much more closely tied up to our common-sense understanding. Strange is the specialist field that delves into these topics with no intention of revising common understanding. Physics clearly has more to gain from Einstein's revision than do our everyday understanding of the world. With regards to our concepts of persons, lives, and survival, on the other hand, there is no such clear special science that will benefit from conceptual clarity. Rather the domain that will benefit the most is the domain of common sense itself. I think advancements like Einstein's can be made in our understanding of persons and survival as well, but this will only happen if we start with what really matters to us about these concepts. Only then can we come up with new ways of thinking that stand a chance of revising the old. Good answers, in turn, have to connect back to the way of thinking under which we cared about the questions to begin with.

1.3.3 The Structure of This Thesis

I have already given some reasons for writing this thesis and some background assumptions required for my arguments to work. The rest of the thesis is divided into six parts, making a total of seven chapters. The main bulk lies in the exposition of my own account and in the comparison of that to other well-known and established ones – this is chapters 3 and 4 respectively. In chapter 5 I consider some potential consequences of my account – were it to be accepted – on ideas and practices that are intimately tied to our concept of personal identity. This will include topics such as the justification of punishment or reward for past

⁶ Ted Sider is a notable example of someone who holds this position. He claims that while questions of causation and personal identity are just about how we think about the world, questions of whether objects have temporal parts whenever they exist and whether every set of objects compose a new object are questions about the structure of the world itself. (Sider 2010) His position will be discussed further in chapter 4.2.2 and 4.2.3.

actions, the force of commitments to future actions, and the retainment of property over time. In chapter 6 I go through some of the objections I expect to provoke and attempt to preemptively ward them off. Chapter 7 is a summary, where I take stock of what I have tried to do, what I think I have accomplished, as well as what remains as topics for further work.

First of all we have to get more familiar with the supposed problem motivating this thesis: In Chapter 2 I will try to explicate the simple view that I suspect most of us hold intuitively (even those of us who have spent much time pondering its flaws). Once this simple and intuitive view is made fairly clear, I advance through a series of steps – starting with a version of physicalism and ending in a mess – to show how this simple view gives a flawed picture of what it means to survive. With this I intend to reveal the need for a better account (and thus the need for a thesis like this). In Chapter 2 I also discuss the different ways of framing the discussion, in part referencing how Eric Olson⁷ partitions the matter – but ultimately also why I will not be following this structure.

At any time (for instance right now!) you can go to the appendix for a lengthy (but hopefully stimulating) read of thought experiments. The appendix is a numerated series of thought experiments sorted both by type and outlandishness. I will at times, in the interest of saving space or maintaining a train of thought, reference the various thought experiments in the appendix, instead of describing them in full within the main text. The series can also be thought of as an expanded and less elliptical basis for the argument I make in chapter 2.2, and finally it can be read as an introduction to the topic, or serve as an “intuition-pump”. My own view on how to interpret these scenarios can be found in chapter 3.3.

In short: Chapter (2) is where I make the case that we have a problem; (3) I try to solve this problem; (4) I compare my solution to well-known solutions; (5) I look at the conceptual and practical consequences of my solution, were we to accept it; (6) I conjure up objections to my view and try to tackle them; (7) I review what progress I have made.

⁷ (Olson 2010)

2 GETTING AT THE PROBLEM

2.1 The Simple Intuitive View

What I want to do here is establish, as a reference, what I think our simple pre-theoretic view of personal identity over time is like. That is, I want to make clear what I take to be the typical pre-theoretic view. This is meant as an explication of an implicitly held view, however, so although I do need the account to be something that most people⁸ would agree with upon reflection, I do not mean to argue that most people express their view in this way. It is also not meant to be the only intuitive or simple conception we have of ourselves. It is, however, meant to capture a conception that is key to understanding how we intuitively think about our own survival. This view is meant to apply to most atheists and will fail to capture the views of those who think we have a soul, or who think there might be an afterlife. The ways in which these views are different, however, will not be important, as the picture I am going to paint for the non-believers will leave them with much of the same difficulties that the believers explicitly commit themselves to.

I take the simple view to be something like the following: A person is a numerically distinct acting and experiencing being that at some time is born into existence. The person endures through a lifetime of acting and experiencing but finally goes out of existence again at death. From the person is born until the person is dead, there is a single subject, a single perspective that gets to experience everything, regardless of how much the person changes during that time. Even if someone undergoes such a radical change that he is no longer recognizable to his friends and family, and even if he himself no longer identifies with his past, it is still thought that the same *experiencer* gets to live this radically different life, rather than it being a case of a new subject coming into existence.⁹

Let us dwell on this point for another paragraph. You anticipate your own future experiences in a way that you do not anticipate anyone else's. In part this could be explained by the fact

⁸ "Most people" is quite inaccurate. In reality I am concerned with an individualistic western mindset, and more specifically with the minds of the people who are likely to read this paper. That there might be cultures or individuals that do not share this view is, however, not detrimental to my argument. I am, after all, going to claim that our naïve conception is fundamentally flawed.

⁹ I mean by "experiencer" an experiencing subject, construed as something that endures through time; transcending memory, but not death.

that you best know how *you* would experience something yourself – better than you know how anyone else would experience something – but that is not the whole story. Even if you imagine that there is someone exactly like you (for whom you would know exactly what the experience would be like), you would still think that *his* future experiences would not be yours, and that you could not anticipate them as you can your own. The key here is the idea of a simple numerical identity of an individual subject – an *experiencer* and an *agent* – that holds throughout every person’s life, no more and no less. This view implies that there will always be a simple answer to the question of who in the future is who (at present) – and that this relation will always only hold one to one.¹⁰

2.2 The Simple View Unravelled

In this part I want to gradually build up a demonstration of the main problem with this intuitive picture. I will start by talking about our concept of death, and through that look at what we would take as possible to survive. In turn, I will try to show how this does not match up with a simple view of person-persistence. Again, the appendix can serve as an extended basis for this line of reasoning.

2.2.1 Death as Irreversible Destruction

Before modern medicine got going, an adequate definition of death might have been something along the lines of cessation of vital functioning, i.e. you are dead when you are no longer breathing and you have no pulse. However, once we got fairly good at resuscitating people from this state, it made more sense to think of this state as a critical condition rather than as constituting death itself. You could survive this state, but only if someone resuscitated you. Once you have the ability to restore normal functioning from some lifeless state, that state no longer becomes synonymous with death. Because of our current medical science you

¹⁰ In general we might say of something that it is identical to several things. We might say of a forest, for instance, that it is identical to a whole lot of trees. What I am arguing here is that we cannot think that one subject is identical to several. Being a subject (at least in the sense that I am after) is all about the numerically distinct perspective you hold. Two subjects cannot be a single subject in this way. In “The Logic of the Trinity” (Bohn 2011), Einar Bøhn advances a view where one thing can be identical to many. While the paper uses the Trinity as an example (and God, the Father, the Son, and the Holy Spirit might arguably be thought of as subjects), his paper does not discuss the identity of these figures specifically *as* subject. My contention is that *as* subjects we cannot conceive of ourselves as ever *being* two subjects. The closest thing we can do, I think, is to conceive of ourselves being the (one) subject of two sets of experiences. As for *thinking* and *acting* as two subjects, I cannot see this working out at all. I will, because of this, assume that this idea of one-many identity does not conflict with my claim.

can now survive for long stretches of time without a pulse and with no measurable brain functioning, as long as you are kept very cold.¹¹ This general development gives us reason to move towards a definition of death more along the lines of irreparability – i.e. you are dead when nobody can fix you anymore, when nothing can restore you to working condition.¹²

Such a conception of death means that if there was a gas leak in my building and I were to suddenly blow up into a million tiny bits, I would then be dead – not because I at the time had no pulse, or no brain activity – but because nobody would be able to piece me back together again. For starters, they would not know what pieces belonged where, but even if they somehow did, they would still not be able to put the pieces back together again. In principle, however, we can see that if some god-like being or technology were somehow able to track all the pieces I consisted of when the explosion occurred, and then somehow could put these pieces back together, just the way they were before the explosion, then this too would actually be something I could survive. There is no apparent reason to think that a person cannot survive the temporary disordering of his constitutive parts, as long as they are put perfectly back together again. It might seem a stretch of the imagination to envision a technology or a being that is able to perform such a feat – it certainly is a far cry from the medical science of today – but all I need for this argument to work is that we agree that it would be possible *in principle*.¹³

2.2.2 Numerically Individuated Parts

If we so imagine that there was this being (or technology) that could put you back together again from a million pieces, the question becomes: would he have to use the very same pieces you were just composed of, or would it be enough to use the same *type* of pieces? More concretely, if he built you up again atom by atom,¹⁴ would he have to put the very same carbon atom back where it used to be, or would any carbon atom do the job? What difference

¹¹ A stunning example of this is the case involving Anna Bågenholm, Norway 1999. (Gilbert et al. 2000). For an excellent overview of how suspended animation is used in current medicine see (Back From the Dead 2010)

¹² The National Conference of Commissioners on Uniform State Laws in 1980 formulated the Uniform Determination of Death Act. It states that: "An individual who has sustained either (1) **irreversible** cessation of circulatory and respiratory functions, or (2) **irreversible** cessation of all functions of the entire brain, including the brain stem is dead. A determination of death must be made in accordance with accepted medical standards." (Jonsen, Veatch, and Walters 1998, 137, my emphasis)

¹³ I mean "in principle" as reflecting of our concepts, i.e. if this somehow *were* possible it would constitute survival and not death. This is what is sometimes called "conceptually possible".

¹⁴ This is an interesting axis to scale in its own right. We could both go smaller and larger – from elementary particles to cells. For the sake of simplicity I will simply use atoms from here on.

could it make whether you ended up with the one or the other carbon atom in a particular spot in your body? As we live our ordinary life, our body constantly swaps out all its molecules – not all at once of course, but what would that matter? I think it would not matter if all my current atoms were suddenly swapped for identical ones – the important thing is how all these atoms behave where they are, and given that all carbon atoms would behave the same (at least with respect to what function they play in a biological organism), that behaviour would not be affected at all. It seems unreasonable to deny the principal possibility of being resurrected in this manner unless we can find the *specific atoms* that the person once consisted of.¹⁵

2.2.3 Multiple Reconstructions

If you have not jumped ship yet, we now agree that a person can survive getting pulverized as long as he is put perfectly back together again, and for this we would not have to use the very same atoms he consisted of; rather it would be enough to recompose him using for each place in his body the same type of atoms that was there before he was pulverized.

The next step then is to ask why we could not recompose more than one of him. Seeing as we do not need to use the same set of numerically distinct atoms he once consisted of, we can now in principle make numerous copies that are just like the person that was pulverized. What we are imagining then is a scenario in which the very same relation that we just agreed would constitute survival in the previous cases now would hold from the one person that was pulverized to several persons afterwards. The upshot of this is that if you were deconstructed and somehow perfectly reconstructed again several times over, you would survive as multiple persons. If this is where you want to draw the line and quit the game, the alternative would be to claim that even though any of these copies *on their own* would have been sufficient for your survival, now that there is more than one copy, none of them are sufficient. How could the second copy ruin the survival of the first? Surely the introduction of such a non-branching restriction cannot be taken seriously as a discovered requirement of our survival.

¹⁵ See appendix A3.4 and A3.5

2.2.4 The Fall of the Simple View

You, the numerically unique experiencing subject, cannot be two places at the same time in the future. The simple view that started us out falls short of making sense of the situation we have ended up with. We have one person, one subject, related to two distinct subjects at a later time. The relation to both these subjects are equivalent to what we think normally holds together a life, and so it stands to reason that what holds together a life is not the persistence of a subject. This means that although it does not reveal itself as such without a bit of exercise, the simple view is never a good analysis of what it means to survive.

2.3 The Many Ways of Framing the Discussion

I have so far chosen to talk about *survival* rather than *personal identity over time*.¹⁶ The vast majority of literature on this general topic talks instead of personal identity, and the question I am to a large part concerned with here is usually thought of as *the persistence question*. Moreover, the persistence question is usually thought to be *the re-identification question*.¹⁷ This means that how this is usually conceived, what we are after is an answer to the question of what it takes for a person at one time to be identical to a person at another time. What are the necessary and sufficient conditions for a person A at time t1 to be identical to person B at time t2. This is the re-identification question.

Eric Olson details how he thinks the personal identity debate should be structured in his Stanford Encyclopaedia article “Personal Identity”.¹⁸ Here he stresses the importance of recognising the persistence question as the re-identification question, and of distinguishing it from other related questions. Interestingly, Olsen is keen to separate the question of persistence from the question of what matters to us about persistence. He criticises Parfit for talking about survival in a way that does not imply persistence, but Parfit does this precisely to allow the investigation to track what matters to us, letting the neat logic of identity sail its own course. In going with Olsen on this we might maintain a neater logic, but in doing so we would be detaching the inquiry from what made us bother with it in the first place. Surely, Parfit has the better end of the stick here.)That is, he had: In his seminal 1984 book “Reasons

¹⁶ This distinction is more or less echoing the distinction made by Parfit (Parfit 1971, 8-9).

¹⁷ (Olson 2010)

¹⁸ (Olson 2010)

and Persons” he surrenders the term, though he still maintains that what matters in survival is not identity – or, then, *survival* per se.)¹⁹

The argumentative structure in this thesis closely resembles Parfit’s work on this topic.²⁰ He is a reductionist about personal identity over time, claiming that it is properly thought of as consisting of other relations such as physical and psychological continuity. He further argues that since personal identity over time simply consists in these other relations, it should be thought to matter to us in virtue of these relations as well. From this premise he uses fission scenarios (see Appendix Chapter 4) to show that these continuity relations in principle can hold one-to-many. As he takes it that identity cannot, he concludes that identity cannot be what matters to us about our survival after all. There have since been many attempts to create accounts of personal identity that allows for identity to hold in these fission cases as well.²¹ All of these come with their own sets of difficulties. For me, however, the true insight in Parfit’s reasoning (or else his fission scenarios) has always been the realisation that the simple view of personal identity is false. Here already we are faced with the problem of accounting for how persons – acting, experiencing persons – are related to what we think of as their pasts and their futures, or more generally what we think of as their lives. Because a person individuated by what he does or what he experiences is properly thought of as a numerically distinct subject, it is unclear how we should think of this subject in relation to a reductionist construal of a life. Finding reductionist persistence conditions for persons is an idle sideshow to this problem.

While Parfit does end up concluding that personal identity is not what matters, he does not get there without setting a stage for this sideshow. To my mind it is not the fact that we cannot construct robust persistence conditions that poses the hard problem, rather it is the fact that we give up the pursuit of a simple non-reductionist account. This difference in our approaches might be explained by the fact that Parfit seems to think that the idea of the subject itself is reducible²², while I on the other hand regard it as a clearly basic idea.²³ Because of this I want

¹⁹ (Parfit 1984, 258)

²⁰ (Parfit 1971, 1984, 2007, 2008)

²¹ Notably (Perry 1972) and (Lewis 1976). More recent accounts include (Sider 2001) and (Moyer 2008).

²² See for example: “We could fully describe our experiences, and the connections between them, without claiming that they are had by a subject of experiences.” (Parfit 1984, 225), or “[T]he existence of a person [...] involve nothing more than the occurrence of interrelated mental and physical events.” (Parfit 1984, 340-341)

²³ Because of my view of metaphysics (Chapter 1.3.2, 4.2.3, 4.3.4 and 6.3), I do not posit a distinct entity out there in the world, when I say of something that it is reducible. I merely claim that the concept is basic to our understanding – that there are fundamental parts of our understanding that hinges on this concept.

to distance myself slightly from Parfit's work, and attempt to build the argument from the ground up myself. In doing so I will be creating what will often seem like an echo of Parfit's. This is not to pretend that I have independently come up with this line of reasoning, but to allow the emphasis to lie with what I think is philosophically most interesting. Building the argument myself will hopefully also allow my perspective to come across without the reader getting caught up in the myriad of slightly different ways of thinking about the topic and how all of these perspectives are related.

Even if we managed to create a definition of personal persistence that could hold one-to-many, the central issue would still remain. Such a reductionist definition would be guaranteed not to describe the persistence of an individual subject – only a non-reductionist account could (Chapter 4.5). As long as we maintain that the acting experiencing person is properly thought of as a numerically distinct subject, there would still be the task of incorporating this perspective into the account.

There is another relatively straightforward way in which a person could be said to persist. We are human beings, and in the biological sense that means we are human animals, or organisms. By thinking about ourselves as animals we have no difficulty in seeing how someone could survive becoming a different person, so long as he is still the same animal.²⁴ But does that mean that we have solved our problem? Well, the fission problem of numerical identity affects this conception as well. It is not clear that an animal cannot survive the temporary scattering of its parts, or the replacement of all its parts with qualitatively identical ones. Additionally, we are also a collection of atoms – and the collection you are composed of now has been part of dying stars and hot dogs, and after you are dead your atoms will be scattered across the world, some of them eventually finding their way back into somebody's dinner. That means that there is a way in which you will become a highly scattered set of atoms, but it does not mean that this scattered set is of any importance to you, or to your life. I will argue that it is like this with the animalist conception as well (Chapter 4.3). It is simply

²⁴ "Animal" is sometimes used in contrast to "human", sometimes as synonymous with "mammal", and sometimes in a more strict biological sense (along the lines of *all living organisms that feed on organic matter, have specialized senses, and typically move voluntarily*). For the notion to be interestingly different from "person" it will have to be individuated by something other than character. Whether your pet dog Sparks is the same animal as dog X might be thought to depend on whether this dog has Spark's memories, personality etc. For the notion of animal to contribute something to this debate, something other than this must be meant. For instance we can think of animal identity as requiring continuity of biological life, or something along those lines. This will be further discussed in Chapter 4.3 on Animalism.

not under this conception of ourselves that we are concerned with questions of personal identity or survival.

3 MY VIEW

3.1 My Life and Me

There are several distinctions I intend to make in this chapter. I think much of the reason why the debate around personal identity is so difficult is that we have several partly overlapping but largely conflated conceptions of persons and their lives. I will start by singling out the conception of a person as a numerically distinct subject of experience and action. This should remind you of what I have been dismissing as the simple view in the previous chapter, but it is when thought of as capturing a persisting entity that this conception does not hold up. As a conception for picking out ourselves and others in the moment, it does not have these problems; on the contrary, the concept is necessary for thinking about persons as subjects of experience and action. I will contrast this conception with one that comes closer to that of a character, the person individuated by his features and properties. These two conceptions of a person I will then hold up to our concept of a life and try to reveal how they all relate.

A person is thought of both as a subject of experience and action, and as a character. The character we can understand as consisting of a collection of properties (the things that make you *who you are*). The subject we can only understand in relation to action or experience (the *you* that experiences this or does that). The subject seems not to be dependent on the character (although the experience is). The idea of a subject seems just to serve as an empty pointer, or a perspective-cursor. You can imagine yourself right now as having different memories, different personality, different friends, work, family, history, but what exactly is this *you* that you imagine would have all these different properties? It is simply the function of imagining it from a first person perspective. The only difference between you imagining *yourself* being this different character and you imagining *there being* such a different character is what perspective you choose to imagine this hypothetical case from. I can imagine being you, but what I am imagining then is not the world being different. Rather I am imagining the world from a different perspective.

3.1.1 I Am a Numerically Distinct Person

At any time I consider my own situation I find that *I am me* and that *I am here*. These claims can hardly be considered empirically vulnerable.²⁵ I am a numerically unique person with a numerically distinct perspective in which I experience and from which I can act. This is quite different from the assumption that nobody else are exactly *like* me, and it is true of me, here and now, regardless of the state of affairs of the rest of the world. We use this way of thinking about others as well: If we have before us a person, we can know that he is in the same way numerically distinct, holding a numerically distinct perspective in which he experiences and from which he can act. Again this is not dependent on whether there are anyone else out there exactly *like* him. Rather, this uniqueness has to do with the fact that each person, individuated this way, constitutes a separate locus of experience and action. Let this idea of a numerically distinct person at a time be signified “person₁”.²⁶

Person₁: *A person individuated by a numerically distinct perspective, experience or action*

This conception of a person we will contrast with the idea of a person as constituted by his various aspects:

Person₂: *A person individuated by person-constitutive properties*

By person-constitutive properties I mean to include memories, preferences, habits, fears, desires, personality, humour, temper, world view – in short, what we might think of as “the things that makes a person *who* he is”. Individuated only by his constitutive properties, the person can be thought to straightforwardly persist over time, simply through there being some person at another time that is exactly *like* him in the relevant ways.²⁷ Conceived of as a person₁ however, my claim is that he is properly thought only to exist *in the considered situation*: he is the one doing what he now does, thinking what he now does, experiencing what he now does etc. On this view a person₁ cannot himself persist, but this does not mean that there is not a meaningful sense in which a person₁ might have a future.

²⁵ That is, if we take “me” and “here” as simple indexicals, not containing any further information.

²⁶ I mean to use “at a time” here the loose way we use the expression in everyday language. I do not mean to signify a literal *point* in time as do for instance Sider (Sider 2001). I do not think the idea of a non-temporally extended person makes much sense. If a person₁ had no temporal extension he could not be thought to be what thinks, feels, acts, experiences etc. – and these are precisely the things I want to ascribe to the person₁. This will be further discussed in Chapter 4.3.

²⁷ The relevant ways being all the aspects of a person important for making him the person that he is.

3.1.2 I Have a Future

If I (considered as a person₁) do not exist at a later time, in what way can I be said to have a future? In what way can *I* do something *tomorrow*? I think the idea of *me having a future* and in particular of *me going to do something in the future* is justified by the fact that I have a deciding influence on some future person₁, and that this influence consists in a power to shape this person by “shaping myself”.²⁸ A future person₁ is in *my* future because what *I* now think, experience and plan for will be retained in a first-person perspective by him. My current intentions concerning myself will be inherited *as is* by this person₁, and so he will see these intentions as being his and about himself. If he is to go to the dentist tomorrow, *I* now stand in the unique position to prepare him for that, and the way I must do that is by preparing myself – as if it was I who had to go to the dentist – because he will inherit the way I map out who these thoughts concern as well. In lack of a straightforward persistence of myself as a person₁, this inheritance-relation constitutes what direct *reach* I have into the future. I can meaningfully deliberate on future actions as well as entertain long-term goals because I have psychological descendants that are moved by these deliberations, that inherit my decisions and reflections and consider them to be their own. But this is not a complete and all-powerful reach. The future actions I now plan for will ultimately be executed (or not) by another person₁. I (perhaps especially²⁹) do not have complete control over this future person₁. If I want to make sure to do the dishes, say, I better do them now, because I cannot with the same certainty make sure that this future person₁ is going to do them; he might postpone the action further, likely recycling whatever rationale *I* used for postponing in the first place. There is no point in just hoping that my psychological descendant will be less of a sloth than me.

A person₁’s anticipation of future *experiences*, on the other hand, is not directly justified by his *reach*. The idea of reach merely describes the influence or power he has on his future, and so while this relation might extend his action-space, it looks like it does nothing in the way of extending his window of experience. This is true, and it is important to realize this difference and not just immediately reach for a new definition. That being said, there is a way in which the person₁’s reach justifies his idea of future experiences that has to do with preparation. In anticipating future experiences, the person₁ that in the future inherits these reflections becomes prepared for those experiences. Also, when considering future actions it is

²⁸ By “shaping myself” I mean simply that this is how I must look at it if I am to in fact shape this future person₁.

²⁹ The degree of control over persons₁ in your future is dependent on what person₂ you are. This is a proper character trait of its own.

impossible to prepare for the action without anticipating the experience of doing it as well. Usually the action relies on external triggers for its timing, and at the very least it will depend on proprioceptive feedback for its execution. This fact justifies a person₁'s anticipation of future experiences when linked to preparation for future action, but what about his expectations of future experiences that are not related to action at all? When you are simply looking forward to some pleasant future experience, a massage, say, what justifies this anticipation? While some story of the structure of motivation could be pursued, part of the justification must on my view ultimately come down to something like empathy. You can rejoice in the pleasure of people you care about and you care about your future self. This means that you should be equally justified in looking forward to the future massage of someone else you care about. While empathy is normally thought to allow us to look forward to or dread the future experiences of somebody else *as if* it were going to happen to us, it on my view becomes part of what explains the anticipation of the future experiences of both ourselves and of others equally.

This way of making sense of ourselves and our futures might seem very cumbersome, and indeed it is. When talking about our future we ordinarily talk of the future person₁ as simply being *ourselves* in the future, and as I point out, this is even the only way we *can* think about him which lets him become prepared by our preparations. The point is, however, that there is an important way in which your future self is *not* you, and that is in the sense that he is a different *person₁* than you. If we fix our minds on this subject-conception of ourselves we are forced to think about our future and our past in this very roundabout way, and in this realization I believe there is great insight.

3.1.3 Person₂-Inheritance and *Reach*

The *reach* a person₁ has to his own future can be thought of as a subject-oriented expression of a more general relation that propagates all his person-constitutive aspects through time. "Reach" denotes a way in which the person₁ has deliberate and intentional power over his future, but the underlying relation enabling this influence is not something separate from that which propagates his person₂-properties over time. An unconscious person is not a person₁, and of course has no reach of his own. He is however propagating the reach of the person₁ that went to bed, along with all the person-constitutive properties that makes him the person₂

he is. The unconscious man is an equally good link in the chain of reach and person₂-inheritance that makes up a life, he is just not a locus of experience or action (not a subject).³⁰

Whether we want to think of this in terms of a causal story or not will depend on what conception of causation we favour. It will depend on what we take to be acceptable causal relata (objects, properties, absences, actual events, agents, mental events etc.) and what we think the relation consists in (dependence, necessitation, counterfactual dependence, transference of energy or momentum, constant conjunction, laws of nature, or whether we think causation is an irreducible basic relation).³¹ I will try to stay clear of the causality debate and rather talk about “where such and such features, intentions, properties or the like *came from*”. Consider a person₁ that is very much the same person₂ as some prior person₁. These person₂-properties might then be *from* the past person₁, or they might be, by some cosmic coincidence, from somewhere else. While I could perhaps express this as a question of whether these are *the same* properties that the earlier person₁ had, this would commit me to a view of numerically unique properties (whatever that means). Expressing it as a question of whether these earlier properties *caused* the later ones commits me to a view of causation that I am uneasy with.³² Expressing it as a question of whether the later properties were dependent on the earlier will require me to qualify this supposed necessity further, as it would certainly neither be straightforwardly conceptual nor physical. I will, then, express it as a question of *origin*, hoping that this will be sufficiently vague and agreeable.

In general we can speculate on whether all purported persistence of numerically unique objects can be reframed in such a way. Usually we divide the world into objects and think of these as simply persisting over time, but perhaps all numerical re-identification claims are properly thought of as reducible to claims about a weighted qualitative identity plus some causal, or at least relational, story. If we pick out features or properties that exist at a time, we can relate these to future instances of the same properties as being either a continuation of them or not. To pursue this tangent further, consider if you will a computer simulation of balls

³⁰ This picture of influence and property-propagation is similar to Shoemaker’s psychological account of personal *identity* over time (His view is discussed briefly in Chapter 4.4).

³¹ For an overview see (Psillos 2002). I am sympathetic to Cartwright’s view (inspired by Anscombe’s) of causation as being an abstraction from perfectly ordinary perceivable phenomena. (Cartwright 1993)

³² I have always favoured actual *forces, events, or actions* as causal relata. For instance, I do not like saying that it was the fragility of the glass that caused it to break, nor the absence of a pillow to soften its fall. Nor do I like to say that it was the glass or its position at some prior point in time that caused it to fall to the floor. Someone with a more liberal view of causation should probably just consider the relation I am discussing here a causal relation.

bumping around in a confined space. The location of a given ball at any given time, indeed the very presence of this ball, can be thought to obtain in force of the ball's history in the simulated world. But as this is a simulation, and not a scenario involving actual balls, the continued presence of a given ball is only realized through the intermediary of the program coping it, as it were, from the preceding state of the simulation. An equally good analysis, then, is that the states of this world are individually generated, though with the use of the previous state plus the rules of the virtual world to determine their content. The virtual balls depend on the rules of the program to "carry them" from one rendered state of the system to the next. But is this really so different from real-world balls relying on the rules of nature? Well, enough of that. I will briefly return to this topic again when discussing what unifies a life in Chapter 3.2.2.

3.1.4 The Idea of a Life₁

The relation extending a person₁'s reach into the future and propagating his person₂-constitutive properties through time can be thought of as being what normally constructs a life. Let us, then, define a technical notion of a life: A life₁ is a single thread of person inheritance or propagation through time. Two persons on such a thread would not have to be identical in any way, just part of the same life-thread, or chain of inheritance.

***Life₁-related:** Two persons₁ are of the same life₁ iff a chain of person₂ inheritance and person₁ influence exists between them.*

I am of the same life₁ as the one working on this thesis tomorrow, because he will inherit everything about me, including all the aspects that make me who I am but also the more direct intentions I have for tomorrow.

Importantly, there is nothing that restricts this life₁-relation to only form a single life₁-path. Although I necessarily am a numerically distinct person₁, I could in principle be sharing my entire past with several duplicate persons₁ sitting somewhere else right now, just as my numerically distinct present experience may come to be part of many lives in the future. Pointing to a person at a time, then, we could be sure we were picking out a single person₁,³³

³³ One could obviously construct scenarios involving visual illusions and the like, where what appeared to be a person was in fact two, or easier still: none. But it would be a perceptual, not a conceptual problem. It would not be a problem of telling what "a person standing there" would mean, but rather a problem of seeing whether this in fact is the case.

but we might in principle be pointing to many lives₁. This person₁ may have many futures, many pasts, or even both.

Mostly what we consider to be lives are indeed lives₁ in this more technical sense. And most of the persons₁ we ever consider are indeed engaged in one and only one of these lives₁. Someone who is fast asleep, or otherwise unconscious, cannot really be counted as a *person*₁ as we have defined him; nonetheless such a person is usually preserving a life₁-strand for future persons₁ to engage in.³⁴ In cases of never-ending coma this is not the case, as there is no person₁ in the future for him to pass the torch to, as it were. In cases of brain trauma where all memories have been wiped, personality has drastically altered etc. it is not the case because although a person₁ might wake up, he will not be on the same life₁-line. The connection from the person₁ before the accident has been severed.³⁵ The person waking up will instead be starting a new life₁, with only insubstantial fragments remaining from the last. This relation should arguably be thought to be one that holds to different *degrees*.

Considering myself as a person₂ on the other hand, I can persist in a more straightforward sense, although only as a *type*. A person₂ is identical to me to the extent that he shares my person-constitutive aspects.³⁶ This identity-relation is also one that holds in different degrees. Normally then, there is a person₁ tomorrow that is of the same life₁ as me and that is the very same person₂ as me. But these are conceptually independent. We can imagine a person₁ that is the same person₂ as me by some cosmic coincidence; that while he is the same person₂ as me, he will not have gotten any of these properties *from* me. (Thought-experiment A5.4 in the appendix describes such a scenario.) Conversely we can imagine a person₁ in the distant future that is connected to me in the right way, making him of the same life₁ as me, but that has become a completely different person₂. A vast life₁ may in principle outlast every property of a person₂ locatable within it. This is the case of the immortals³⁷ – they live long lives₁, gradually changing their personality, memories, abilities etc. until they have become

³⁴ In addition we might consider lots of things going on in sleep that will have a bearing on the person₁ waking up the next day. Beyond the obvious aspect of dreaming, the sleeping person might undergo physiological changes that will have an impact. In general, then, the sleeping person can be considered more than a mere vessel, part of the evolution of the life he is connected to – but at the bare minimum he is such a vessel: a keeper of life-strands.

³⁵ This is also to some extent true of severely demented persons₁ – they do not have an effective reach into the future, although the cases are usually not so black and white. A lot of memories could be gone, for instance, while the personality might still linger on and even continue to adapt to new experiences.

³⁶ Here the better candidate is one of similarity over a weighted set of relevant central aspects, rather than just general unrestricted similarity.

³⁷ As discussed by Parfit here: (Parfit 1971, 24), and here: (Parfit 1984, 304)

persons₂ with no overlap in their person-constitutive properties with the person₂ they used to be (and this they do many times over). What ties the persons₁ locatable on this life₁ together is this chain of inheritance I have been discussing. We allow for this change in person-constitutive properties in our everyday notion of a life as well. The child grows up to become quite a different man; all the while, there are no days that were not heavily influenced by the previous. But what if we imagine two contemporary persons₁ that are the exact same person₂, how would they relate to each other's futures? To answer this we must look closer at what is granted in having a future.

3.1.5 I Want to Survive

I want to survive; and as I have already extensively discussed, this wish, in its primitive form, can be described as wanting the person₁ that I am now to exist in the future. But as I have also extensively discussed, this wish cannot be granted; it is a flawed idea grounded in a failure to separate different perspectives, different ways of considering ourselves. I cannot exist tomorrow in this way, so I will have to settle for the next best thing – but what is that? I have already presented what I think are useful elements in thinking about this: Having reach to a future person₁ and there being a future person₁ that is essentially the same person₂ as me. But how do they interact?

I think we are naturally inclined to care about our personal future lives₁ in particular, allowing for great change in person₂.³⁸ But this notion misses slightly on a more reasonable idea of personal survival. It is flawed in that what the reach grants are possibilities of *action* and *influence* in future parts of the thread, and still it fixates on the numerically distinct person₁ having this reach. Since the reach does not extend the person₁'s experience, then surely, if you have nothing to add or subtract from what the other person₁ would do in this situation, you have nothing to gain from being the one whose life₁ goes on. If there existed another person₁ that was identical to you as a person₂, surely he could carry on your life-story just as well as you could. If I am right, another notion of survival is called for. The need for this can further be made salient by considering the scenarios A5.1 and A5.2 (in the appendix) where even though the current person₁ has no *reach* into the future, he should nonetheless be considered

³⁸ Especially those specifically addressed by this thesis (who assumed they survive as persisting subjects), should plausibly be inclined to care more for life₁ relatedness than person₂-persistence.

to survive since what is vital to him, as a person₂, is preserved. I suggest the following analysis of personal survival:

Personal Survival: *A person₁ survives for as long as his life-story is carried on through life₁-lines, where a life-story can pass on from one life₁-line to another only where the life₁-lines contain essentially the same person₂.*

A life-story can be thought to contain a single *developing character*. This is different from the fixed character implied in “person₂”, and more like the character in a fictional story – one with a *character arc*. When we point to a person₁ that has a fission event in his future, we are pointing to two life-stories (or simply “two lives”, as I described it in the introduction), but are we then pointing to *two* such developing characters? I think we should avoid this by saying that the person₁ we are pointing to is the character in two ongoing life stories. The two ongoing stories is of course not visible as *two* in the given situation. Only from a reclined perspective, for instance in retrospect or in imagining, can we see this situation as being part of two life-stories.

At this point the discussion might seem to be heading into familiar tracks: “Is developing character A at time t₁ the same as developing character B at time t₂? What if there is a fission event?”. It would seem that all the same problems crop up again, only this time with developing characters instead of persons. Fortunately, however, I am not postulating new entities in the world every time I employ a different concept. What I am saying is that we could, and indeed we do, describe a life story as being the life-story of a character, or even a person; we just have to remember what this means, or rather what it does not mean. The idea is not a perfect one, and it might crumble under load, but it is not a load bearing pylon in the analysis I am advancing. I will at times use the idea of a developing character, but I will, for clarity’s sake, refrain from using “person” this way.

Regard the following illustration of an ordinary life:

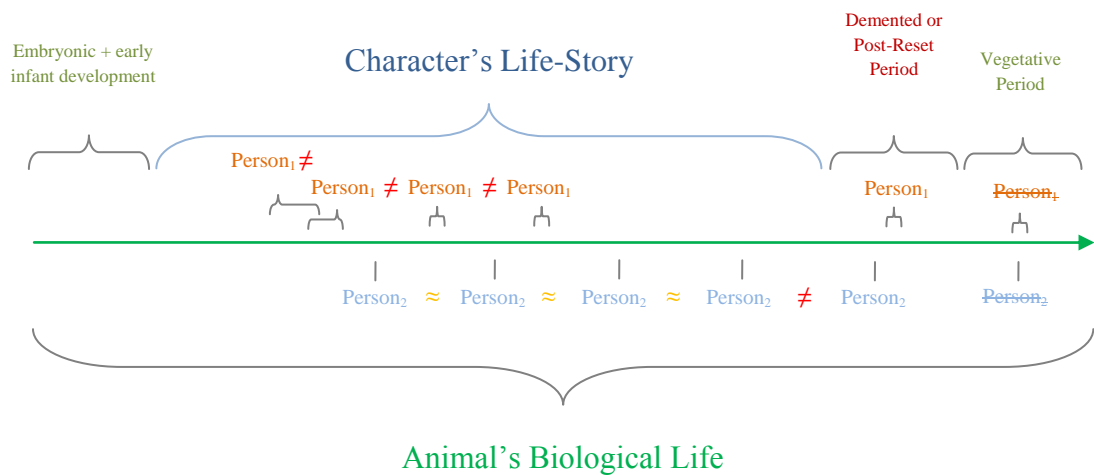


Figure 1: Ordinary survival

On my view a any person₁ locatable on this line survives for what I have demarcated the “Character’s Life-Story”. From a folk-psychological conception of ourselves this is the story that makes the most sense to follow. Ordinarily this life-story simply is the human animal’s progression through its biological life, with the exception of pre- and potential post-character (or even other-character) periods, it is however picked out by different criteria. In the next illustration the life-story outlasts the animal it starts with:

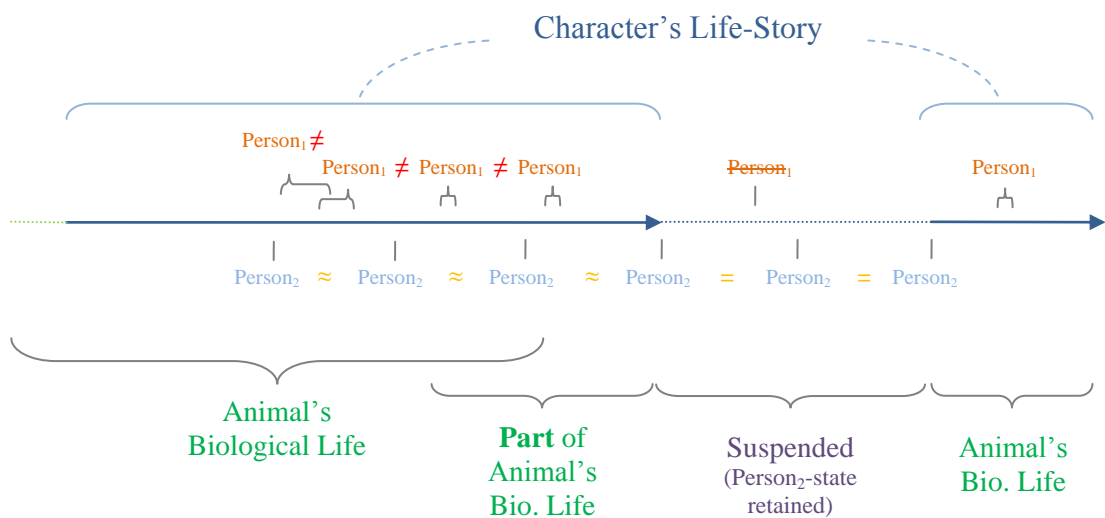


Figure 2: Exotic but possible survival

Figure 2 shows a life-story that spans three distinct animals. Looking at the animals involved, we first have one dying a natural death, its composition gradually imposed on another creature. This second creature has its original composition gradually replaced, then after living for a while gets pulverized, its final state recorded and stored. Later the third animal is artificially synthesized from the ground up, its composition dictated by the stored record of the second creature's final state. This third animal goes on living until its natural death.

My argument here is that following the character makes more sense than limiting our gaze to any single one of the creatures involved. The first creature does make sense on its own, whereas the second and especially the third do not. While we could opt for telling the story of three distinct creatures we would be equally justified in tracking the character instead, and this is what you would do if you knew him. I think the only reason we would want to track the animal's biological life instead of the character's life-story is if we thought it held the equivalent of a soul – that it somehow retains the same numerically distinct subject. If we throw out this idea of an enduring soul-like subject-substance, there is little reason to treat this life-story differently to an ordinary life. Indeed, it is my contention that this is what it ultimately comes down to.

It is worth remembering that we are free to go beyond the borders of the life-story to see where it comes from and what it leads to. You can hold that you are a person, and that an embryo or even an early infant is not a person, and still you could say you were once a baby. This is just like saying that a certain collection of soda cans used to be a car. If a car has been scrapped and its metal used to make soda cans, we might say of these soda cans that they “were once a car”. This is not meant to imply that there were soda cans around *at that time, in the form of a car*. Rather it is meant to imply that what was once a car, now has become a collection of soda cans.³⁹ We can see something becoming something else and still track properties from the first thing to the second. A dead body is something different from a living person. The set of molecules that is released into the atmosphere and reintegrated in plants and microbes upon cremation is something different from a dead body as well. Still, we can explain the composition of the dead body, or the amount of CO₂ released into the atmosphere upon cremation, by appealing to properties that the *person* once had.

³⁹ There is of course a difference between these examples in that the way an embryo becomes a person is far less arbitrary than how a particular car becomes a collection of soda cans. The point is, however, that we can track something beyond what we specify it as.

Consider again the second animal in Figure 2. If we consider him in a situation some time after he is integrated into the depicted life-story, he is a person₁ that is life₁-connected to the past of a different animal. We can thus say of this animal that it used to be part of a different life-story, but also that he used to be a different animal; it depends on what conception we use to pick out his past, *what past* we are thinking about.⁴⁰

3.1.6 Survival and What We in the End Care About

On the account I have been arguing for, my own survival becomes a commensurable good – its value given by the value of my roles and projects.⁴¹ This also means I might not care as much about it as I did about the prospect of enduring as a subject. Since there is no further fact securing any future experiences to *me*, the current experiencing subject, I might choose to only identify with and worry about those future beings that are the continuation of my psychology, or only those that are the continuation of my organism, or my body, or only beings that are just the way that I am now, or no one at all! The fact that somebody in the future remembers thinking these thoughts, remembers concluding that he was not going to have a future, say, does not prove these earlier thoughts wrong. In practice, however, this does put a certain pressure towards thinking that I am going to be the person who ends up with my thoughts. At any rate the picture is more complex now. I naturally care about my future life₁, but perhaps I care more about some specific part of it, some project I am engaged in perhaps, or something broader, like the survival of my group, my family, my religion, my ethnicity, my nation. The collective memory of a group, its gradual adaptation to stimuli, its power over its own future state – it has a life₁ much like the person₁. Its members could also use this perspective as their own – see themselves as a cog in a great machine, but identify with the machine as a whole, not just the part they control.⁴² In doing so they would be no more mistaken than the one limiting his perceived identity to his personal survival. Personal survival is, however, as good as it gets on my view, if what you really wanted was persistence as a subject.

⁴⁰ This is the same sort of reasoning applied by Sider in his "Stage-View", discussed in Chapter 4.2.

⁴¹ I should perhaps amend this statement somewhat. What I mean to say is not that my life has no intrinsic value. Rather it has no special intrinsic value, relative to any other life. What special value it has can no longer be thought to come simply from it being *my* life.

⁴² List and Pettit explores just this sort of possibility in Chapter 9 (especially 9.3) of their recent book "Group Agency" (List and Pettit 2011).

3.2 Interpreting the Thought Experiments

It is time to look at some of the thought experiments from the appendix and see how my account fares in dealing with them. I will run through the five categories that the appendix details, not necessarily covering each particular scenario but rather covering the types of challenges that each of the categories represent. In general there will be two sorts of challenges to survival recurring throughout these stories: One is the loss (from the world) of some central feature or property of the person considered. The other is a loss of connection, the loss of influence or reach, if you will, of the current agent with regards to the future. On its own, both types of loss weaken the degree of survival, but they do so in quite different ways.

3.2.1 Surviving Accidents

In these scenarios an unfortunate person suffers various injuries. On my account the person₁ waking up in the hospital is never strictly identical to the one that suffered the injury. What matters is how the two are related. There are two important measures: One is how well the one waking up in hospital maintains the central features and properties of the one that suffered the injury. The other is how connected the two are, in the sense of whether the actual similarities are *from* the person that suffered the injury and whether her reach is preserved. In the first version of the scenario where the woman loses her legs, the woman waking up in hospital is almost the person₂ as the woman suffering the injury,⁴³ and furthermore she gets all these person₂- properties *from* this woman. On my view (and I would think on most others') she constitutes the survival of the injured woman to the fullest. In the later versions of the scenario she still represents the best candidate for the injured woman's survival, but the degree of survival is considerably lowered. The person₁ waking up in the hospital in A1.2 is disconnected from the past life₁ by not remembering any personal memory. She is connected still to this life₁ by retaining much of the other aspects that make the woman involved in the car crash the person₂ she is. Not only does she have the same personality, the same habits, the same interests, etc., but these traits are from the person that fell asleep at the wheel. A large part of the person₂ is lost (her memories), but a large part is also retained (her personality, her preferences). Her pre-accident reach into this future will also be severely impeded. In A1.3,

⁴³ Given, of course, that her psychological identity, for instance, is not heavily defined by how she uses her legs. (If she was all about her jumping, for instance, this would be hard to adapt to, and it would consequently constitute a larger hit to her person₂-makeup.)

the person₁ waking up in hospital is disconnected from the past life₁ by having lost most of the aspects that made her the person₂ she was. She is connected still to this life by retaining most of the memories of the person₁ crashing the car, but these memories no longer have the meaning they used to, and so in a very real way they are not quite the same memories.

I regard the loss of personality in A1.3 as constituting an even harder blow to the woman's survival than the loss of memory in A1.2. Such destruction of personality is not only seen in exotic trauma cases (like the much cited Phineas Gage incident⁴⁴) but in the not so uncommon cases of frontotemporal dementia (FTD).⁴⁵ To appreciate the practical importance of our conception of survival we need look no further than to cases of FTD. "I'm walking next to her and going all the way down to hell", the husband of a FTD-patient states in an article about the illness.⁴⁶ While we will always care about what happens with the immediate remains of a person, it makes a big difference whether or not we believe that the late stage FTD-patient is still the same person we once knew. Consider the prospect of taking care of someone who looks exactly like someone you loved, but that does not show any affection towards you, and who's personality you do not at all recognise. Why someone would endure this torment is more understandable if he thinks that it is still the person he loved experiencing this radically different existence. In my view there is no enduring experiencer. In addition, the late stage dementia patient does not maintain the life-story of anyone, or at least she does not do this well. She is not the same person₂ she used to be, and this new character is not shaped by the past intentions, goals and experiences of any person₁. She is still the same organism, still the same animal, but that cannot be what matters. She is disconnected from the life₁ that this biological animal used to be engaged in.⁴⁷

3.2.2 Suspended Life

Here are some cases where the person's life is conserved in unusual ways. This poses the question of what sorts of processes can support a life. When I talk about connection or reach I am intentionally not specifying any sort of causal structure. I do not think it matters how the

⁴⁴ (Harlow 1848)

⁴⁵ University of California, San Francisco's "Memory and Aging Center": <http://memory.ucsf.edu/ftd/overview>

⁴⁶ (Chen 2010, 37)

⁴⁷ Importantly, this is not to say that we should not care for the late-stage FTD patient, we obviously should. What I am claiming is that the husband in this case has little reason to care for *this* person in particular. (I should qualify this statement further, in saying that I have understood the typical late-stage FTD case to be like I have described. If however the reality was that the old personality regularly shined through, as it were, then the case would be much more complex.)

details of the story goes as long as there are persons₁ who inherit their person-constitutive features from earlier persons₁. When Kim is reduced to a cloud of particles in A2.3, the set of molecules that he just consisted of no longer retains his life. The body that would normally have granted him reach into the future is lost. But his reach is still not lost – rather it is secured through another intermediary: It is retained in the mind of the mysterious traveller as a perfect representation of his complete physical composition. Because the mysterious traveller goes on to assemble a person from this representation, the person that was pulverized is given, in the perfectly ordinary sense of the term, a future – though the way in which this ordinary relation is established is anything but ordinary. If it turned out that God was destroying the universe every other second, but that he always went on to recreate it just the way it was before he destroyed it, that would make no difference for anyone's survival.

In the case of long-term cryo-stasis (A2.2) another aspect of a person's life is made salient – his relations to the outside world. If we imagine a person waking up after being kept in such a state for a thousand years, the world will likely be very different.⁴⁸ If we imagine that everything he knew and cared about was long gone, there is a very real question of whether there remained much of his life at all. Though there would still be the memories, we can imagine that many of the central aspects of this person₂ might have been lost.⁴⁹ To what extent this scenario would pose these problems depends on what the central aspects of this person's life was. If all this person cared about was shooting heroine and lying on his back in the grass contemplating God, and there was still heroine and grass around, the many years that had passed would present little challenge to this person's survival. If he on the other hand was heavily invested in some project that no trace could be found of, or generally if his psychological identity was strongly tied up to his job, his friends, his role in some group or society at large, and these things were all gone, then the person₁ waking up could perhaps not be said to be the same person₂. But would he be of the same life₁? Although the person₁ entering the cryo-chamber in some sense has a perfect connection to the person₁ waking up, we might conclude that he has a quite limited reach to him after all. The preparedness, the motivation, the goals and ambitions, all its content might have become vacuous. If the world

⁴⁸ Stated in a most cumbersome way: A person₁ enters a cryo-stasis, his internal features and properties are retained by his frozen body for a thousand years, then this body is thawed, a person₁ wakes up, the question being to what extent this is the same person₂ that was frozen a thousand years ago and to what extent this person₁ is engaged in the same life₁ as the person₁ that entered the cryo-chamber.

⁴⁹ I am unclear about exactly how much external factors should be involved in defining a person₂. This is however just a pragmatic question where if I change this factor I would simply have to redescribe some of the other parts of the account.

has become so different that the reach from the previous person no longer makes much sense, then we might interpret this as a broken life₁-thread. On the other hand, we might say that his reach was perfect because he really was in a position to prepare him for this future, had he only known what it would be like. Whether we call this survival or not, if all he cared about and all he defined himself by was lost, this future would be of little value to the person₁ entering the freezer.

3.2.3 Replacing Parts

The commonplace scenario of A3.1 describes what we would ordinarily call a life.⁵⁰ It also fits the technical notion of a life₁, at least a healthy part of it does. For any person₁ within that life₁, that person₁ has a firm grip on his near future, and while he might over time become quite a different person₂,⁵¹ this happens without breaking the chain of influence: There is no disconnect of reach or inheritance, only an evolution of character.

In A3.2 we have an example of organ replacement by outside intervention. This means that if a substantial part of the body is replaced this way it might at some point not count as the same animal. As for the person's survival as I have sketched it, these interventions are of little consequence and especially if the new body parts work exactly like the old ones. In A3.3 on the other hand the brain is swapped out with another brain that is very different. Although the new brain has the same genetic formula and is of comparable age, its structure is shaped by the life of the clone, not the customer. And so it is the clone's life₁ that is continued by the person₁ waking up after the replacement.

In so far as what happens in A3.4 makes any sense (if this god is correctly described as having done anything at all), it might constitute the end of both the animal and the body that was, but it makes no difference at all for the life of the person involved.

A3.5 and A3.6 continues in pretty much the same vein. These are examples of life₁-lines perfectly maintained through exotic intermediaries. What sets them apart from scenarios like A2.3 is simply that all the matter is swapped out from one stage of the life to the next. I am not sure how atoms are properly thought to persist through time; whether they are best

⁵⁰ Some four-dimensionalists would call it a person! Lewis, for instance, takes the person to be the whole of a four-dimensional worm; what is present in any given situation is on his view not a person, only a temporal part of him. (Lewis 1976) More on this in chapter 4.2

⁵¹ This is especially true if he lives a very long life.

thought to endure as numerically distinct objects or if they perhaps are better thought to constantly cause their own successor (it is not apparent to me either what the difference would be between these two). For my view of persons and lives, however, this question is irrelevant. Since we have argued that a change of atoms, carbon for carbon, does not affect the function of the person, the change of atoms does not change the strength of the $life_1$ relation either. What determines the structure of the new set of atoms is the old set of atoms much like in an ordinary life, except here there is a period with no change whatsoever. The fixed state that is being transferred is not a $person_1$ – not primarily because it is not in the form of a material body, but because it is a fixed state that cannot act or experience, much like the cryo-stasis phase in A2.2. It does however contain all the non-relational aspects of the $person_2$ that was captured, including the earlier $person_1$'s intentions, preparations, and plans, granting him reach to the person that later is materialized.

3.2.4 Fission

The fission scenarios are the ones that really start to poke at the sensibility of our intuitive ideas of life-long subjects. In these stories a single $person_1$ has more than one future $life_1$. The first case is the classic Hemisphere Split à la Parfit⁵² in A4.1. This example is most useful for probing the conception of those that prefer more materially grounded identity claims, as well as more down to earth thought experiments. We shall return to something of the kind when discussing animalism in the next chapter, but for now we will zoom ahead to the neater, cleaner, fission scenarios of full-blown material duplication.

For all the following fission scenarios the following diagram (Figure 3) will be applicable where O is a $person_1$ that has a fission event in his future. In the farther future there are two distinct $persons_1$ (A and B) that are both related to O by a $life_1$ -strand, but that are not themselves related to each other in this way. They are not in each other's past or future, rather they are contemporary $persons_1$ with a common past.

⁵² (Parfit 1971, 4-). Parfit in turn cites Wiggins as its originator.

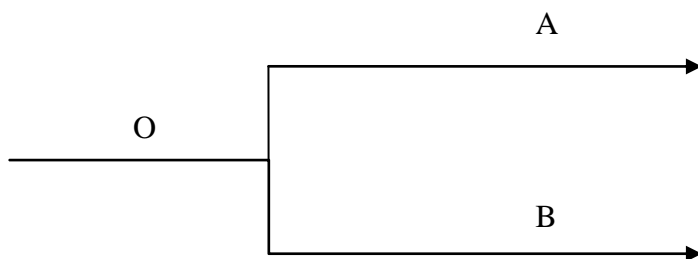


Figure 3: Fission

When A and B meet in A4.2, some time has passed since the fission, and so in addition to their common past they each have a more recent unique past as well. The two lives might have taken different turns, meaning A and B might be quite different persons₂, if not, they might be pretty much the same person₂. They will both think that what they remember as their childhood is *their* past, and they will both be right in thinking so. What has changed is that now there are two persons₁ engaged in two separate lives₁ that both have this one childhood in their past. On my view, neither A nor B *is* O, but they are *both* O's successor. The way they are related is the same way that any person₁ is related to another person₁ of the same life₁. O has two futures. In none of these futures is there a person₁ who is two persons₁. However, there will be, at least to begin with, two persons₁ that are the same person₂, meaning that if one of these persons₁ would perish, the person₂ would still not be lost.⁵³

In A4.3 and A4.4 (as well as A5.2) the question is raised whether a short-lived post-fission life₁-strand is a bad thing, and for whom it might be bad. The pre-fission person's survival is not at stake here, but what about the post-fission person₁ that finds himself on the short strand? This is the person₁ that finds himself "still in the outbox" after the teleportation machine has already materialized his copy somewhere else. He has no reach to this other future, he cannot influence it in any way, all he has to deal with, so to speak, is his immanent destruction. Depending on how he views the situation this might be a painful experience, and that is obviously a bad thing. But beyond that, is this really a bad deal? Does he have good reason to despair, or would despair be a confused reaction to what is happening? I would say he is confused to despair; which is why I have gone so far as to include the relation he has to this other future as part of what we should call survival. The current part of the other life₁-strand contains an identical person₂ to him. Everything he wants to do, everything that defines

⁵³ When I say that the person₁ would perish, I should really be saying that this person₁ would not be life₁-connected to any future person.

him as a character, every goal, preference, motivation, it is all doubly present and by that other instance secured a future. This is as good as it gets. The fact that he cannot presently add or subtract anything is offset by the fact that he has nothing to add or subtract to the choices of an identical person₂. The only thing he can in fact add is a desire to not be stuck in the machine he finds himself in, but this is one of the few differences between the two: The other person₁ is *not* stuck in a machine. He does not need to fight for his freedom, he does not need to think his way out of a machine, he is already out, getting on with his life. The person₁ stuck in the outbox does not have to deal with anything, he does not have a life₁ ahead of him, or at least not one lasting more than a few seconds. He should be content to survive through the other life₁. There is no important difference between this situation and the one described in A5.1. The person₁ survives by a future that he (so narrowly construed) is disconnected from. As long as nothing important happens that would have, had he been connected to it, changed this future for the better, then all is well. The psychological stress of this situation, if misinterpreted, is of course a bad thing regardless of what the right interpretation may be. To gauge how bad we think this stress is on its own (disconnected from the badness of the perceived situation), we can imagine putting someone in this situation, making him think these things, but then letting him out, telling him it was all a bad joke. Actually it is better than this, because in these scenarios no life is tainted by the trauma of experiencing such a horrible joke. Still it is somewhat bad to subject someone to such a fission scenario. On the other hand, if the person is prepared for this situation and does not freak out, the “Error Checking and Correction” teleportation of A4.4 is a far safer way to travel than non-ECC teleportation. My short answer, then, is that if a potential traveller thinks he can psychologically handle the brief situation of being disconnected from the future he will survive to, he has every reason to prefer this sort of delayed-disintegration ECC teleportation over ordinary teleportation.

In A4.5 we are presented with a person₁ that is faced with a fission event in his near future. One of the post-fission life₁-threads holds the future he wants, a future of professional surfing. The other holds a future as a soldier in a war, a future he does not want at all. While the one waking up with a rifle in hand has no reach into the surfing future, the contemporary person₁ that *does* is the same person₂ as him, and so as with the previous scenarios he should be content with having this other person₁ holding the reins to his continued life-story. Everything about him continues, and while he is in no position to influence this future, a person *exactly* like him *is*. He is in the end superfluous to the life he cares about. From this he could choose

either to end this superfluous life₁-branch, or to embrace the future he did not really want and start preparing for combat. If he chooses the latter he will in effect be having it both ways at the same time. He will survive as both a surfer and a soldier. But he only has reach into the soldier future and so this is the future he must face up to and prepare for. The decision depends on whether the life he is disconnected from is what in the end matters to him, or whether he wants to try his hand at this other future as well. While it is psychologically hard, I do not think he has any more reason to go on with the soldier future now than the pre-fission person had of wanting this soldier-copy to be made in the first place. As for the option of “swapping futures” by shooting the surfer and taking his board, this will in effect differ from shooting himself only by the addition of a traumatic memory to the surfing future, one that might very well ruin the fun of surfing.

In A4.6 we are faced with some related moral questions: Is it permissible to create copies of yourself and send them to bad places? Are the copies expendable, and can this be decided by the pre-fission person₁ or would we have to make the copy first, and then ask *him*?

In as much as we ever have the right to decide what to do with our future, and I think we think that we do, we have the right to send off copies of ourselves to bad places. They of course have the right to object and overrule as far as that is possible, but this is no different from an ordinary case of a person changing his mind about a former promise. It make little difference if astronauts on their way to the moon suddenly get home-sick – they have no choice but to finish the journey they started out. And this situation we allow the astronaut to get himself into. At the same time we would not tolerate someone sending out copies of himself left and right to frivolous ends, thinking he will not suffer these futures just as much as he will enjoy his future here on earth. On my view the person₁ about to be copied is wrong to think that he will experience the one future but not the other. If he does not realize this we might deem him unfit to decide over these additional futures, just as we normally might decide that someone that does not act prudentially are unfit to decide their future. In general we do not think it is right to allow people who does not realize the weight of their commitments to bind themselves to detrimental futures, be they children, demented, or intoxicated. If someone persuaded by the view I have been arguing decides that he only cares that he has *a* happy future, he might try to make multiple copies of himself in order to increase his chances. But if this simultaneously increases the chances for most of these lives to be bad then we should perhaps not think that this personal proliferation is morally permissible.

There is also something to be considered in the psychological toll that goes with preparing for the one way trip in A4.6. Because the preparations are done before the fission, this burden will also have to be carried in the much longer life₁-strand here on earth. Therefore, although he should empathize with the traveller, he should not focus too much on this strand being a dead end.⁵⁴ The person₁ at the dead end does on my view survive through his duplicate back on earth. He would in a sense be less wrong to imagine that he comes home to his family, say, than seeing this dead end as being his death, even though there will be no event corresponding to this “coming home”.

3.2.5 Beyond

These scenarios are generally more outlandish or else difficult to place within the previous categories. I have called upon some of them already, but the scenarios of A5.5 and A5.6 presents us with a question of proximity: Does my proposed relation of survival apply across great distances in space, or even backwards in time? The notion of survival I suggest allows for a person₁ to survive through the future life₁ of another person₁ without them being themselves of the same life₁, as long as this other person₁ is the same person₂ as him. If there is an identical person₂, not by his side, but across the universe, can he still be thought to survive through the future of this person? While I think the knowledge of such a duplicate might offer some comfort with regards to closure of projects (like wanting to finish a book or complete a thesis), there is still the much more pressing matter of all those you leave behind in this world that are hurt. The existence of their unbereft counterparts will be of little solace to them. This issue can be more easily probed by considering a duplication of our entire galaxy at the time someone is killed here on earth, but in the duplicate galaxy this killing is averted. The killing here on earth is still wrong because of all the people that will suffer a loss in this world. Their actual suffering is not offset by their non-suffering duplicates. These duplicates do however grant these sufferers a second path of survival, and so if the original galaxy was annihilated in the process, then that *would* on my view counteract the killing.⁵⁵ Then all these people would survive *only* through the duplicate galaxy. As for A5.6, it is not clear what we should think. It will probably depend on whether we hold an eternalist or a

⁵⁴ I am assuming that the mission is a short one; that there will still in all essence be a person₂ identity between the strands when the one-way mission draws to a close.

⁵⁵ This would once again be like God annihilating the universe (this time the galaxy) and then recreating it just the way it was; except in this case there is a slight difference in the copy (the absence of the killing).

presentist position on what exists.⁵⁶ My view offers no immediate interpretation, but then this is a very strange scenario.

In A5.7 I think we easily dread the future of this creature as being our own future, but I also think we are mistaken to do so. While the creature might be the same animal as you, and largely the same set of molecules, you have no reach into its future and it will not retain any of your projects or concerns, it will not remember you, and it will not remember anyone you care about either. While the creation and treatment of any of these creatures is detestable, your friends and family will have limited reason to seek out *this* creature in particular. Two ways in which they would have special concern for this creature is that it is what has become of you, and it superficially resembles you. Just as we care about what happens to the dead body of someone we knew, we would care about what happens to this creature in particular. And just as we would react more strongly to a counterfeit video of someone being tortured if the one being tortured was made to look like someone we cared about, we would care what happens to this creature in particular. In A5.8 the perception of “remains” is somewhat weakened by changing form twice over, but if culturally trained to we are liable to care about such things as the ashes from a cremation, so this might still be a factor. Whether the other factor kicks in depends on whether the generic sufferer is made to look like the former person or not. On inspection, then, these two scenarios are much more comparable than they might seem at first glance. Here the animalist (Chapter 4.4) might give an account that better coincide with our *prima facie* interpretation; but only through conflating the idea of a biologically persisting animal with the idea of a persisting subject will we be impressed with these results. More about this in the next chapter.

A5.9: In this scenario the usefulness of an account like mine becomes pressingly clear. It is really only hell-bent metaphysicists that would claim that the proper thing to track here leaves us with two people with highly fragmented life-stories, remembering only every other year of *their* lives, while falsely remembering the interluding years that was really lived by their

⁵⁶ This is the question about whether all that exists is what *presently* exists, or whether what exists also include everything that ever *did* exist as well as everything that ever *will* exist. (Zimmerman 2008a) and (Smart 2008) defend opposing answers to this question in (Sider, Hawthorne, and Zimmerman 2008, Chapter 5.1 and Chapter 5.2) The reason I think this question might be related to how we should interpret this thought experiment is that if you do not grant the present any special status it might be easier think that it does not matter that the rest of your life story lies in the past, instead of the future.

brother.⁵⁷ On my view the past and the future of one of these persons₁ is whatever he is life₁-connected to. It does not matter that these lives span over different bodies. That is, it does not matter for the question of where his past and future lies. It does of course matter for each of their lives that they share bodies. If one of the brothers wants to be physically fit, while the other only wants to sit in and smoke cigarettes and eat pizza, this will amount to no small point of contention between them.

A5.10: The first person₂ gradually disappears, while the second gradually comes to life. The slowness of the transition is irrelevant. If the person at the end of the transition has all the properties of the second person₂, then the second life₁ will be continuing from here on, and this period will have constituted the slow demise of the first life₁. In this whole period there will for any experience or action undertaken be a person₁ that is the subject of this, and he will be a mix of the first person₂ and the second person₂.⁵⁸ If we let the things experienced and done in this period remain with the person after the transition (in the way I imagine it now), then he will end up as mainly the second person₂ with a splash of the first person₂'s insights, outlook, perspective etc, along with memories from this period of having done things that now feel foreign to him.

3.3 A Grand Finale

At the end of this chapter I have chosen to include an attempt at a unifying thought-experiment. Its purpose is of course to make it easier to understand what my view states, and I really do think it can be quite clarifying. It does, however, introduce a lot of information at once, and so if you find it more confusing than clarifying, feel free to skip ahead to the next chapter.

⁵⁷ For a clear example of someone who would argue this, see: (Thomson 2008) Also, in order to make the following point, I have used the term “metaphysicist” rather than the more common “metaphysician”, because I do not see the logic of the latter term. It is, after all, not a term for a practitioner of *meta-medicine*. (I guess one could argue that “physician” is a strange term for a medical doctor in the first place.)

⁵⁸ i.e. a person₁ in this period will have a mix of the properties that makes up the two different persons₂.

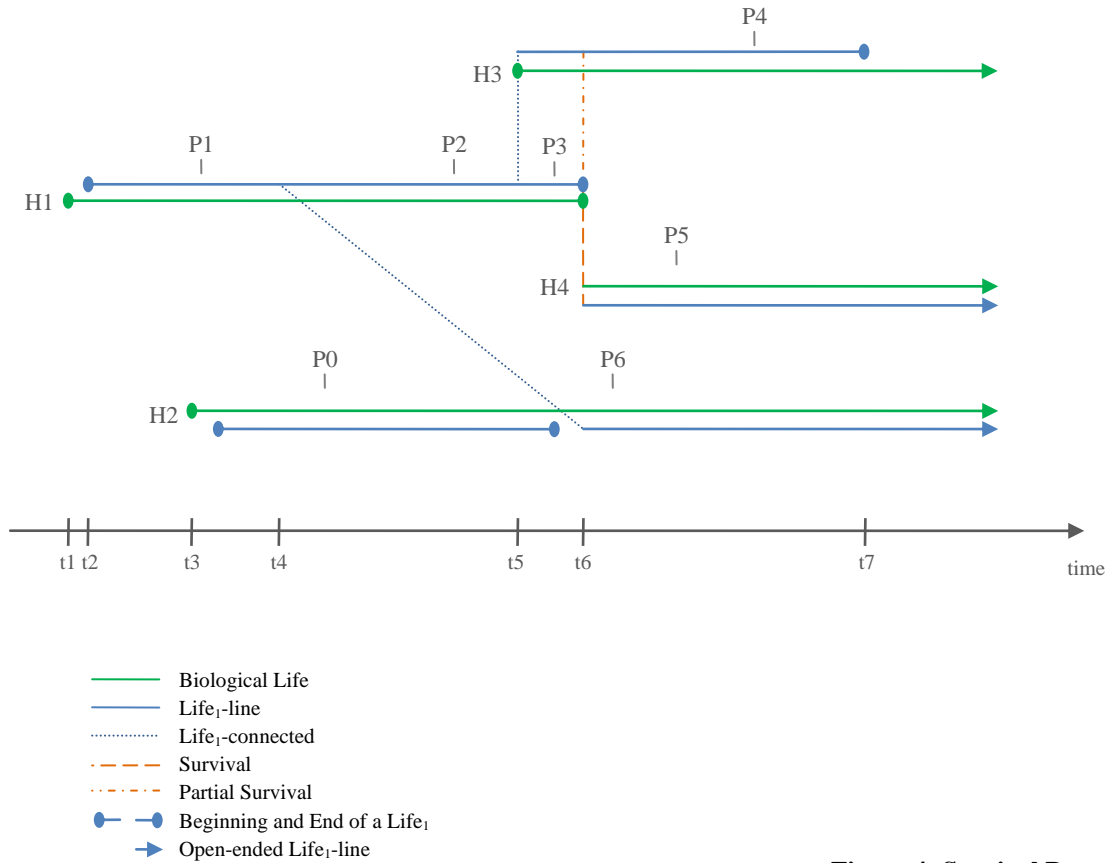


Figure 4: Survival Bonanza

At t_1 a human being (H1) is created (conceived, if you like). Around t_2 , as H1 is coming to be a functioning person, a life₁-line begins. At t_3 another human being (H2) is created. At t_4 a perfect scan is made of H1. The scanned structure is gradually imposed on H2 until he, at t_6 , finally becomes physically qualitatively identical to H1@ t_4 ; this also means that he has all the non-relational person₂-properties of H1@ t_4 . At this point H2 is also life₁-related to H1 before t_4 ; he is no longer life₁-related to his animal past (H2 *after* t_6 is not life₁-related to H2 *before* t_6). At t_5 a physical duplicate (H3) of H1@ t_5 is created. At that time H3 and H1 are person₂-identical. At t_6 , H1 perishes; this death also signifies the end of a life₁-line. The person₁ (P3) who finds himself on this life₁-line just before t_6 is still not doomed though. For one, he survives partially through H3 who at that time is very much the same person₂ as him (not much time has passed since t_5 , and nothing important or dramatic has happened in either life₁-line). Secondly he survives *fully* through H4, a being who by sheer chance springs into existence at t_6 and who is the exact same person₂ as P3. In order to imagine that every relational property relevant for this specific person₂ is retained, we can think of H1's death to consist in him being vaporized, and that H4 is materialized right where H1 is vaporized. This way, even all the *relational* person₂-properties that makes P3 the way he is, is retained.

Where there is a life₁-connection, either through ordinary passage or through fixed states, there is automatically survival. Where I have marked “survival” or “partial survival” there is ordinary or partial survival *in spite of a lacking life₁-connection*. Also: In general, the lines in this diagram are meant to start and stop more gradually than I have been able to draw. For instance, I do not mean to imply that there is a *point* in which a human being is created, or in which he dies.

The life₁-line that starts at **t2** branches twice, both at **t3** and **t5**, and so the person₁ (P1) located on this line around **t3** is strictly part of three lives₁. He can meaningfully prepare for the situations P2, P3, P4, P5, and P6 will find themselves in, but of all of these it is P6’s situation that is closest to him. P6 (which is H2 a while after the forced rearrangement of his physiology⁵⁹) is further ahead in time than for instance P2, but he is closer by life₁-line. This is so because between **t4** and **t6**, the life₁-connection between them is simply mediated by a frozen state, conserving a person₂.

Importantly, the partial survival connection at **t6** does not make it the case that P4 has a branching past. There is partial survival connections constantly running between these life₁-lines as long as they contain, in essence, the same person₂.⁶⁰ The point is that as long as they are not different enough to influence their future lives differently, they can adopt each other’s life-stories. P3 is strictly speaking not in P4’s past, and P4 is strictly not in P3’s future. However, everything P3 cares about in his life (his roles, relations, projects, and intentions) is carried on by P4.

P0 is not life₁-connected to P6, nor does he survive through him. His only connection to P6 is that he is the same animal as him, or at least he is physically continuous with him. P0’s future gradually comes to an end as his influence is erased by the forced recomposition. While we could say of P5 that he cannot remember P3’s situation because the memories are not *from* P3, there is another way in which he perfectly remembers – his memories are just like they would have been, had they been from P3.

⁵⁹ Olson might arguably call this a *different animal* (I am in fact inclined to myself) but this is not important here. I call him the same animal, not because I think this conveys any deep truth about animalhood, but because of the continuity of constitutive parts during the rearrangement process. If we call him a different animal we will simply have to call him the same “something else” – *mere body*, perhaps.

⁶⁰ The one line drawn is thus just an example – it is most relevant where it is, however, because this is the end of the life₁-line he is on.

4 COMPETING ACCOUNTS

4.1 Parfit

4.1.1 Parfit's Argument

Parfit argues that personal identity is not what matters for our prudential concern, i.e. our concern that we have especially for our own future. That might seem contradictory, and it might be the wrong way to phrase it. Prudential concern picks out the sense in which you arguably can anticipate your own pain at the dentist's tomorrow in a way that you cannot anticipate the pain of someone else who also has a dentist appointment at that time – that is our intuition at any rate. Most people assume that this difference just comes down to the fact that you would be identical to the person having the pain in the first scenario, but not the second. Parfit, however, argues that this cannot be it; personal identity cannot be what accounts for this difference. His polemic follows two main strands, the first one is a reductionism about importance (which he in a later paper restates as “realism about importance”), and the second argument is the argument from fission. The latter is the one that has received the most attention. I will briefly survey both strands of his argument:

Reductionism about importance

Reductionism about importance is the idea that, simply put, when something that conceptually consists of something else is considered important, it is only so in virtue of the considered importance of these constituents. Human life is only important because it contains conscious people enjoying themselves, say, and even though we say we value human life, if we could imagine a scenario containing human life but void of conscious persons having a good time, we would rescind this appraisal of *human life simpliciter*. Parfit argues that personal identity just consists in certain other facts – and consequently that it must be these other facts that matter (morally and rationally), not personal identity itself. In his 2007 paper, Parfit remodels this argument into what he calls “realism about importance”.⁶¹ Here the notion of non-mattering higher level concepts are defined as cases where, relative to the facts at some lower level, some fact at a higher level is merely conceptual. Given a grouping of trees, a copse is

⁶¹ (Parfit 2007)

merely conceptual.⁶² We learn nothing new about the world when we hear that there is a copse where we already knew there was a group of trees, what we learn is rather about the term “copse” itself.⁶³

Fission

The other line of argument is the argument from fission. Here Parfit imagines a double brain-transplant where each of his hemispheres would be placed in its own body. The intuition is that if just one hemisphere survives, that person would be him. If both of them survive we cannot say that they are both *him* as they would obviously be *two* numerically distinct persons. Still we can see no good reason for the pre-fission person to restrict his prudential concern to the cases where only one of the hemispheres survive. Thus, Parfit concludes, what matters for prudential concern cannot be identity.

4.1.2 There is Identity, but it's not What Matters

Parfit does discuss straightforward numerical identity in the re-identification sense as well. His criteria for this is that there is psychological continuity, that this continuity holds in virtue of the persistence of a substantial part of the person's brain, and finally that there is no fission.⁶⁴ As opposed to most other contributions within the personal identity debate, Parfit does not try so hard to make this definition work. Rather he is out to make the point that it can't be something like this that we are after. The criteria are meant to count as sufficient, not necessary, and they are meant to serve as a widely acceptable framework for the discussion. As opposed to how I framed the matter in Chapter 3, Parfit does allow that persons normally survive through straightforward numerical identity across time. And the sufficient criteria for this are the ones I stated above.⁶⁵

What I argue is that the relation of simple identity that seems so important to us in fact *never obtains*. What I take Parfit to say is that while persons normally persist, this is *never what matters*. Now, the way in which Parfit takes persons to normally persist is *not* the naïve idea I

⁶² (Parfit 2007, 7)

⁶³ Parfit is careful to point out that although this is what we learn, this is not what the statement was about. The statement was about there being a copse, i.e. a group of trees, at some location.

⁶⁴ (Parfit 2008, 177, 189)

⁶⁵ Parfit presents a neat table listing various scenarios and comparing what different views of personal identity (including his own) predicate of these, in (Parfit 2008, 182)

appeal to, but a reductionist account of psychophysical continuity. My view does not in this way brake with Parfit's. I think however that his line of argument will have lead philosophers to pursue a sidetrack – namely that of construing an identity relation that will be applicable even in the fission cases, so that identity can be what matters again.

4.1.3 Gendler – Identity is What Matters After All

In “Personal Identity and Thought-Experiments” Tamar Szabó Gendler argues that even though Parfit's fission scenario shows that identity and rational prudential concern come apart, this is not because prudential concern is not about identity.⁶⁶ He argues that it *is* identity we care about after all. Identity plays a central conceptual role in our world view and is not replaceable by any relation weaker than identity. What explains our rational prudential concern in the exotic cases that do not strictly contain identity, is that they emulate cases of identity so closely that we are rational to mistake them for it.

Gendler argues that identity is what matters after all. He does this by pointing to how the concept of identity is situated at the centre of a network of related and important concepts like desert and commitments, and that the psychological continuity relation that Parfit advances to replace it no longer evokes any of these connected and highly important concepts. He concludes that it really is identity we care about after all, but that any individual considering Parfit's exotic scenarios would be rational to count them as containing what matters. Not because they do (as they don't) but because the envisioned scenarios so closely resemble real cases of persistence. We are rational to take them to contain what matters, then, because we are rational to use coarse rules of thumb in the first place, rules of thumb that are unable to distinguish these deceptive scenarios from the real deal. With this account he means to have reconciled the intuitive idea that identity is what matters with the intuition that what matters is contained in these scenarios.

The solution comes from conceding that we are in fact being fooled in the latter. Should we be content with this account? I think only in so far as we are satisfied that we are still being fooled by these examples even after being told how they work.⁶⁷ Alternatively we could also

⁶⁶ (Gendler 2002)

⁶⁷ That is not a tall order for perceptual illusions, but those are created far upstream the perceptual interpretation chain. This is a conceptual conflict, what we want is a conceptual understanding that makes sense on careful inspection, not an explanation that says we cannot understand it even if we look really close.

be content with this explanation if what we were asking was how it can be rational to perceive the thought-experiment situations to contain what matters, i.e. why it would be rational to apply an attitude concerned with identity to cases that actually do not involve identity. How it can be rational to use coarse rules of thumb and so be fooled by clever deceptions, as it were. I think neither of these hold true. We want to know what to say about these situations when we carefully examine them. His solution in the end claims that while the non-branching clause is understandably suspect it is nonetheless correct. While the structure of Gendler's argument is philosophically quite interesting, he does not manage to restore identity as what matters in survival. The non-branching clause still makes for an untenable position.

4.2 Four-Dimensionalism

4.2.1 Four-Dimensional Metaphysics

Four-dimensionalism is, simply put, the view that objects are four-dimensionally extended and made up of both temporal and spatial parts. Specifically the view claims that for any time an object exists, the object has a temporal part at that time. An object with both spatial and temporal extension can be described as a four-dimensional *worm*.⁶⁸ The sum of all its spatial parts that are located at any given point in time is called the temporal *slice* of the worm. Conversely an instantaneous object (it only has spatial extension) that may or may not be part of a given worm is referred to as a *stage*. The stage is thus not defined in terms of a larger object, whereas the slice is. Ted Sider discusses two accounts of personal persistence warranted within the metaphysical framework of four-dimensionalism.⁶⁹

The Worm-View

We can at any time see ourselves as consisting of many spatial parts (arms and legs). Just as properly, the worm-theorist argues, we can regard our complete self as consisting of many temporal parts. Ourselves last year, ourselves right now, and ourselves the coming hour – these are all different temporal parts of the same person. A person is never fully present at any

⁶⁸ The image of a “worm” can be motivated by imagining a two-dimensional disc that exists for some time. This object considered in a four-dimensional framework amounts to a three-dimensional cylinder where the length of the cylinder corresponds to the disc's temporal extension. (With an object that is extended in three spatial dimensions (like a ball) the four-dimensional shape becomes harder to imagine as its temporal dimension must be drawn along a fourth dimension. The logic remains the same, however, so the imagery can still be helpful)

⁶⁹ (Sider 2001) The following summary will be based on Sider's discussion on pages 152-208 in this book.

time on this view – only a temporal part of him is. The worm-view solution to Parfit’s fission problem is that the whole life up to the point of the fission is part of *two persons*. The rest of these two persons are to be found after the fission, but here their parts are distinct – non-overlapping. There is then, simply two persons who share a common part, the part before the fission occurs.⁷⁰ What ties the temporal parts of a person together is a separate matter. In “Survival and Identity”⁷¹ Lewis suggests that this person-unifying relation is the very same relation of psychological continuity that defenders of the psychological approach (See Chapter 4.4) take to establish personal identity. For Lewis, however, this relation is not a relation of identity (the temporal parts of a person are not identical), rather it is the relation that holds uniquely between all the different temporal parts that are properly ascribed to a single person.

The Stage-View

Sider advocates an alternative to the worm-view described above. He calls this alternative the *stage-view*. There is nothing different in the ontology of the stage-theorist, Sider points out, there is just a contention over what our ordinary terms pick out in this ontology. On the stage-view a person isn’t a four-dimensional worm, but rather an instantaneous stage.⁷² To account for what we pick out as identity over time, Sider appropriates a counterpart-theoretic model.⁷³ On this model our ordinary terms pick out instantaneous stages, and what is picked out as the continuation of a particular stage depends on the concept by which we are picking it out. We can consider the standard example of a statue/lump of clay.⁷⁴ The idea here is that there is a perceived problem of there being two co-located objects where we want to say there is only one. The statue is not the same as the lump of clay it is made of, it is reasoned, because the two have different persistence conditions. The lump of clay can survive being squashed, whereas the statue cannot. Sider’s account tells us that while the object picked out by “the

⁷⁰ This is the view argued for by (Lewis 1976). Perry gives a slightly different view that seems to count three persons/lives, namely the two overlapping persons and the third which encompasses both post-fission branches. (Perry 1972)

⁷¹ (Lewis 1976)

⁷² The stages are on Sider’s view instantaneous, just like the slices, but unlike slices they are not defined relative to a temporally extended worm. (The term “stage” dates back to at least Shoemaker’s account in (Shoemaker 1984), where “stage” is simply used to talk about fission-cases without assuming either personal identity or non-identity at the outset.)

⁷³ This is properly associated with (Lewis 1968) although in this context the counterpart relation will hold across time, rather than across possible worlds.

⁷⁴ The example originates from (Wiggins 1967)

lump” and the object picked out by “the statue” are one and the same instantaneous stage, what object we later identify as being the continuation of this statue/lump is dependent on whether we perceive it as a statue or as a lump of clay. If we consider the object as a statue, we will track the shape over time, whereas if we consider it as a lump of clay we will track the clay, regardless of shape. A person is the continuation of an earlier person, then, iff the two stands in the appropriate counterpart relation, i.e. as “person-counterparts”. For Parfit’s fission problem the solution is that both of the later persons are related in the right way to the past person, however they are not related this way to each other. This allows us to say that both of them are counterparts of the pre-fission person but not of each other. Sider reformulates this to say that they *were* identical to the pre-fission stage. Notice here that as a person is an instantaneous stage, it is never the case that a person *is* identical to a person at another time. Sider uses “*was/will be* identical” to imply a counterpart relation holding between the two. Your future counterpart is the one you *will be* identical to in the future. You are not strictly identical to your counterparts.

4.2.2 Issues

The immediate problem with the worm-view is that it does not recognize persons as being present at any one time, rather what is present is merely a temporal part of the person. The person itself, on the other hand, is an object that covers the entire lifespan. A person on this view is what I think we would ordinarily call a *life*. In addition to this, since the pre-fission worm-part is part of *two* persons, pointing to what we would ordinarily take to be one person before the fission would either have to be interpreted as pointing to *both* these persons or else as indeterminate between the two.

The stage-view fares better in this respect, but it too has its flaws. Sider himself mentions the objection that persons on his view are too brief (zero temporal extension) to be ascribed the different things we normally ascribe to persons. In particular he looks at beliefs, saying one might object that beliefs take time. This objection he repels by pointing out that while beliefs require certain relations over time, it can nonetheless be instantiated *at* a time.⁷⁵ But surely this is the repelling of a straw man objection. While *beliefs* arguably can be thought to be had by persons *at times* (in the way we might think *properties* are), Sider says nothing about acting, thinking or experiencing. How can any of these fit into a single point in time?

⁷⁵ (Sider 2001, 197-198)

While I have several minor issues with Sider's account, my main dissatisfaction lies in the framing of the problem itself. This likely stems from the fact that we are not primarily after the same thing. This is what I will discuss next.

4.2.3 A Different Objective

Sider's main concern is that of revealing the basic structure of the world itself.⁷⁶ This is quite different from my objective. He sees the border between questions about the world itself and questions merely about how we think about the world as running somewhere close to the border of basic ontology. Questions of causation or of personal identity are on his view just questions about our way of thinking about the world, whereas he takes questions of whether there exist temporal parts or whether every set of objects always compose new objects to be questions about the world itself.⁷⁷ His main goal is to establish four dimensionalism as the best account regarding these deep questions, and the question of personal identity then becomes the question of what in this real-world account our natural language terms might pick out when we talk of persons. In light of this I find it odd that the candidate he advances is a non-temporally extended slice defined in a frame of reference taking into account Einstein's special theory of relativity. My guess would be that our natural language terms grasp something a bit more fuzzy than that.⁷⁸ This highlights how different our projects really are: while I want to hone in on a kind of object that best satisfies our folk-psychological predicates, Sider is after an object that best fits his real-world ontology project.

With his thought experiments, Parfit demonstrated how our intuitions become conflicted when we envision scenarios where we can no longer, even logically, attribute our survival to a simple persisting substance. That is what is so illuminating about these thought-experiments – they show us that whatever we might tell ourselves that we believe about our not being souls and about the composite nature of our mind, we still hold the naïve idea of ourselves as these simple enduring substances. That is why these thought experiments are so provoking. Accounts that seek to dissolve the logical inconsistencies of these scenarios through elaborate

⁷⁶ (Sider 2001, xiv) Sider also has an upcoming book where this real-world ontology project will be addressed more explicitly. (Sider 2012)

⁷⁷ (Sider 2010)

⁷⁸ To be specific: I would think that our natural language terms do not grasp something so narrowly defined in a completely unintuitive framework. (I am not saying that our natural language terms would be fussy in any framework.)

metaphysical frameworks are not attacking what I take to be the more interesting problem, and may even serve to cover this problem up.

4.3 Animalism

Animalism (or the biological approach, as Olson originally called it) is the view that we in essence are human animals,⁷⁹ and that psychological continuity is neither necessary nor sufficient for a human animal to persist through time.⁸⁰ The view has gotten some traction through writers such as Paul Snowdon and Eric Olson.⁸¹ So how could this be a controversial view? We are definitely animals. The tension can be made apparent by considering examples such as brain transplants or part replacement. Thought experiments A3.3 to A3.6, as well as A5.8, all highlight this issue. Animalism holds you to your biological life – even when your person-constituting features go elsewhere.

No doubt there is a way in which we are animals and our persistence can be described biologically as the life of an organism, the problem with this view, however, is that it is not as biological continuers that we care about having a future.

4.3.1 Borrowed Significance

We can see that the concept of a biological organism fits us, but it is not through perceiving ourselves in this way that we care about our persistence. Biology tells us that we are a particular species, and this perspective holds great explanatory power, but it does not imply that if we were to evolve into a different species over the next hundred thousand years we would have lost everything. As for the life of an individual organism, there are many organisms that follow a developmental path that should make us think twice about using biological entities as our starting point. The Pando, a clonal quaking aspen is what appears to be a great old forest of trees, but because of its biology it in fact counts as a single six-million-kilogram organism, potentially a million years old.⁸² Another prime example is the "Immortal

⁷⁹ (Snowdon 2003, 47). While Olson argues that animalism is compatible with us only contingently being animals, it is only conceivably so if *animals* are only contingently animals (animalism is in other words not open to a view where we are contingently animals but essentially something else); to this Olson also adds that he is *inclined* to think that we are essentially animals. (Olson 2003, 320)

⁸⁰ (Olson 1997, 17-18)

⁸¹ Especially Olsen's book (Olson 1997) is heavily cited.

⁸² (Mitton and Grant 1996)

Jellyfish”, *Turritopsis*. It is a creature that is able to revert back to its infant polyp-state after reaching full maturity, meaning it could have several lives, or life-cycles. Biologically however these will all count as a single persisting organism.⁸³ We could imagine an organism like the Pando that housed many separately functioning minds, or better: An organism like the immortal jellyfish that had a developing personality and that stored memories, but only within each life-cycle. It would matter little that the organism survives the entering of the next cycle if everything that was not already coded in its genome was wiped clean. This would be comparable to the prospect of surviving through a biological clone of yourself when you die.⁸⁴ Such a clone would ordinarily not count as the same organism since the cloning was not something the organism did itself. If on the other hand our biology enabled this (rather than our technology), biologists would characterise this clone as the same organism, and the animalists would have to follow suit. That is, they would either have to follow suit, or account for how they could back out.

4.3.2 The Somatic Approach as describing a *Way of Being in the World*

The somatic approach is a view of personal identity closely related to animalism. It is held by among others J.J. Thomson.⁸⁵ This view claims that we simply are our bodies. Again this is in some sense obviously right. As with the animalist conception, however, the somatic approach is contentious in purporting to be an account of our *persistence*. On this account, where I am in the future, is wherever my body is. This view might not be vulnerable to the idea of a self-cloning organism (as that might count as a string of different bodies pertaining to the same organism). It is however vulnerable to the same thought experiments of brain transplants and body parts replacement that I mentioned above (A3.3 to A3.6, as well as A5.8). The idea that I could survive with the head of another person, but not with his body (because the main bulk of the body is in the “body”) seems like an obvious non-starter. I suspect the view might get some traction with those who worry about being brain-body dualists. The kind of motivation I am alluding to here is the push towards identifying with our *whole* body, as opposed to just our cerebral aspect. There is certainly a sense in which our embodiment is essential to what

⁸³ (Piraino et al. 1996)

⁸⁴ “Biological clone” as opposed to “physical clone” which is what I have previously been talking about. A biological clone is like a monozygotic (“identical”) twin, or the famous even-toed ungulate *Dolly*. A biological clone shares the complete genome of the organism in question, but nothing beyond that – i.e. none of the variation that comes from the environment, none of the personal development.

⁸⁵ (Thomson 2008)

we are. This is not the sense produced by a biological or material construal of our body however. Rather it is the sense, once again, dictated by folk psychology.

Our folk psychology is tied up to how we operate in the world, not how that in turns is mechanically realised. The body can be construed as essential to a person only as a way of being in the world. If you woke up and discovered you had been living in a dream world, and that your real body did not fit the biological description of a human being, that would not touch your idea of yourself as a *person* unless your real body performed radically different from the virtual body you were used to. Imagine discovering that you were really created by a god, not through biological evolution, or that you were really made of silicone, not of carbon like you thought. These discoveries would not force you to reconsider your idea of yourself as being a person – and you would not have to consider yourself dead either.

On the other hand you would be hard pressed to imagine waking up from this world and discovering that you were something like the internet. Being something without localized senses and a localized possibility of action would have to be quite unlike what it is like to be you. To be a full-fledged person you would arguably have to operate in the world by some body by which you could act and sense. But this embodiment is not the same thing as essentially being your body *as described by a biologist or a physician*. Being embodied creatures in a world is arguably just as fundamental to us being persons as any other aspect determined by our folk-psychology, but the scientific account of how these features work cannot serve as our starting point. To the extent that we can be said to essentially be our bodies, then, “body” has to be taken as a way of being in the world, not as denoting a particular set of molecules or a biologically specified function or machine (i.e. organism).

4.3.3 Stuffy or Functional Animals and Bodies

The previous paragraph relates back to a question about how we should think of the animal appealed to by animalists. While biology might define an animal (or an organism) in a way that would permit it to survive having its brain, or even its head replaced, it is unclear what biology would have to say about the scenario of teleporting animals. The question is to what extent an animal can replace its constituting stuff, and by what means. I could see biology going both ways on this, thinking that the animal arriving on the one end of the teleportation was a new animal or the same that entered on the other end. It might seem that going the one route or the other, animalism collapses to either the somatic account or to something close to

the psychological approach (Chapter 4.4). To the extent that animalism has something new to offer it is that the animal per definition has to persist by way of its ordinary biological functioning. And this is indeed what Olson claims: “[...] one survives just in case one's purely animal functions—metabolism, the capacity to breathe and circulate one's blood, and the like—continue.”⁸⁶ To understand just what he means by this it is helpful to consider what he says about detached heads and brainstems. While Olson thinks a detached head on life-support might still be the same animal,⁸⁷ a replacement of its brainstem would be its death. According to Olson you could not survive the replacement of your brainstem, as it is the organ that is “chiefly responsible for your life-sustaining functions”⁸⁸. He uses a thought experiment much like A3.4 to consider what would happen if a demon removed your brainstem for only a thousandth of a second and replaced it with a qualitatively identical one. His answer is that you would cease to exist and a new animal would come to be:

So for a thousandth of a second there is no self-directing event that coordinates the activities of your parts in the unique way that biological lives do. For a thousandth of a second there is no living organism there, but only a corpse so fresh that its heart is still beating. (Olson 1997, 142)

This seems to imply that if an animal had only a short lapse of function of his brainstem it would not survive; at least if this was corrected from the outside and not by the animal itself. At any rate there would be no way to resurrect Kim in thought-experiment A2.3, if this is the right account.

4.3.4 A Different Objective Again

Zimmerman makes the point that one of Olson's arguments commits him to denying the existence of any other composite objects but animals.⁸⁹ The argument in question says (roughly) that there is an animal where you are, and that this animal is thinking; then, since you are the one thinking, you are that animal.⁹⁰ Zimmerman points out that you can substitute “an animal” with any other reasonable object that can be said to be where you are, and to be thinking. Olson's reply is that, although he is very unhappy with it, his solution is to deny the

⁸⁶ (Olson 1997, 17)

⁸⁷ (Olson 1997, 133-135) This also means that he should give a different answer than Thomson to the body-swapping twins scenario (A5.9)

⁸⁸ (Olson 1997, 141)

⁸⁹ (Zimmerman 2008b)

⁹⁰ Paraphrased from “the thinking-animal argument” (Olson 2003, 325)

existence of mere bodies, hunks of matter, brains, heads, and anything else that might be a rival candidate to what is thinking. Olson then goes on to point out that this problem is not his alone: “Any sensible account of what we are faces its own version of the rival-candidates problem.”⁹¹ It is clear from this discussion that what Olson is trying to accomplish with his account is something very different from what I am trying to do in this thesis. As it was with four-dimensionalism, the goal seems yet again to be to uncover some underlying reality of what objects there really are. The approach is to first find what particular object we are and then, by revealing this particular object’s nature, seeing what persistence conditions we truly have.

To illustrate the mindset that goes into this sort of approach we can consider what Sider says in the introduction to *Contemporary Debates in Metaphysics*.⁹² He is talking about what philosophy can contribute, in the face of scientific (or more generally: empirical) findings about the world:

The particular-property pattern keeps recurring. It appears that every fact about the world boils down to particulars having properties. So it would seem that the world contains two different sorts of entities: particulars and properties. We have already uncovered a general fact about the world.[...] And we did it without detailed input from the scientists. (Sider, Hawthorne, and Zimmerman 2008, 2)

But surely this is not a deep insight into how the world is? It is an insight into how we structure our thoughts about it. For me, “animal” and “person” are not the same concept, but they are both perfectly applicable to me. The conception of myself as a person reflects more of what matters to me, than do the conception of myself as a biological organism. What I am *independently of the ways we can think about me* is simply not something we can define. At least I cannot see how that would make sense.

4.4 The Psychological Approach

The psychological approach to personal identity over time is generally viewed as hailing from Locke’s *Essay of Human Understanding*. In his *Essay*, Locke clearly divides the persisting person from the animal. He does this by tying personal identity over time to unity of

⁹¹ (Olson 2008, 39)

⁹² (Sider, Hawthorne, and Zimmerman 2008)

consciousness (constituted by memory), and animal identity to unity of life (biological life).⁹³ For Locke, your personal identity is given by your memories. This means that if two human beings were to swap all their memories they would have swapped personal identity as well. This is the case of the prince and the cobbler.⁹⁴

Locke's view faced some difficulties however. First the memory connection was argued to be circular: It is arguably part of the meaning of "memory" that the one that remembers an experience is the one that experienced it. Another objection was that the account breaks one of the axioms of identity, *transitivity* ($A=B$ and $B=C$ only if $A=C$). This is because a person A can experience something remembered by person B ($A=B$), and person C can remember something experienced by B ($B=C$), and still person C might conceivably not remember anything experienced by person A ($A \neq C$).⁹⁵

In "Personal Identity: A Materialist's Account", Sydney Shoemaker advances a view that meets these challenges.⁹⁶ In his account, "memory" is replaced with "quasi-memory" (memory defined specifically without the presumption of identity between the person experiencing and the person remembering). Furthermore, the idea of a simple memory *connection* is replaced by the idea of continuity through a *chain* of connections ($A=C$ iff there is a chain of memory connections between them). On this account, then, what makes a person A at time t1, identical with a person B at time t2, is that there exists a continuous chain of strong q-memory connections between them. However, this definition is still faced with the problem of fission (Appendix Chapter 4). In response to this, Shoemaker incorporates a non-branching clause. This means that person A (at time t1) and person B at (time t2) are identical only if there is no branching in the q-memory chains (as there would be in cases of fission). This has the counterintuitive implication that you will cease to exist if someone makes a perfect copy of you.

It is from this sort of psychological approach, faced with the option of a non-branching clause, that Parfit concludes that identity cannot be what matters for survival. (Chapter 4.1)

⁹³ An Essay Concerning Human Understanding, Book II, Chapter XXVII, §9 and §12 (Locke and Woolhouse 2004, 302,304)

⁹⁴ Essay, Book II, Chapter XXVII, §15-16 (Locke and Woolhouse 2004, 306-307), I will return briefly to this example in Chapter 5.1.3.

⁹⁵ These are approximately the objections of Bishop Butler and Thomas Reid, respectively – discussed in (Perry 1975, 99-118)

⁹⁶ (Shoemaker 1984, 80-88)

Shoemaker does boarder on this conclusion himself, saying what we care about in surviving is not lost in cases of fission, though we would no longer exist.⁹⁷

4.5 The Soul Theory and Other Non-Reductive Views

What about just thinking of ourselves as immaterial indivisible souls? Would that not free us from all the issues I have been discussing? In a way it would – but not without committing us to a lot of other problems. The idea of a soul as a separate non-physical substance gives us a way to free the experiencing and acting subject from the subject’s memories and personality – both of which we know to be physically realised in the person’s body (viz. brain). This conception also supports the possibility of reincarnation and general life after death without any restraints on what that life must be like. Furthermore it denies the possibility of fission. In the double-hemisphere transplant (Appendix 4.1) the soul theorist would either have to say that the soul ends up in just one of the bodies after the surgery, or else that it ends up in none of them.⁹⁸ If we should think the soul ended up in just one of them, the body would either have to be thought impossible to resuscitate, or that a new soul would form, or else that it could go on biologically functioning with no experiencer inside – like one of Chalmers’s zombies.⁹⁹ Equally for the scenarios of teleportation (A3.6), the question becomes whether the soul could be transported along with the teleportation signal. More importantly, however, such a view commits us to an untenable range of problems that has to do with how this immaterial substance can causally interact with the material body. This is the problem of substance-dualism. I am not going to go into these issues here, but I take their metaphysical complications to far outweigh any simplicity that might be gained in this context.¹⁰⁰

But perhaps we could imagine there being a similar simple and enduring *physical* substance within us that holds our subjective perspective. Perhaps it could be thought that the physical brain, the brain cells working in unison, gives rise to something else (perhaps something like a field) that should properly be thought of as a new and *separately existing substance*. That this substance is the actual subject of all experience and action normally attributed to the brain/person/animal. And that this substance endures as a numerically separate entity for the

⁹⁷ (Shoemaker 1984, 119-121)

⁹⁸ Unless he should think that souls could be multiply instantiated, but that would mean that souls were like properties, not substances; and this would render the whole point of souls moot.

⁹⁹ (Chalmers 1996, 93-)

¹⁰⁰ For more on the substance dualism and its shortcomings see for instance (Kim 2006, 29-52)

animal's entire lifespan, not by continually being replaced with ones qualitatively identical and continuous with it, but actually enduring, numerically identical, throughout the person's life. This then might be what is lost in the teleportation scenario.

Insofar as a model like this can be made intelligible, this picture is still not justified by the data. In fact it is not a falsifiable theory at all, so what support it could have in empirical data would be mere conjectures. What conjectures we can make, however, points us in the other direction. Brain processes are highly distributed; this we know from cases where one or more fail without affecting the rest. The fact that a person who suffers such a partial failure is often unable to tell that something is wrong, is highly indicative of a distributed functioning of the mind as well.¹⁰¹

¹⁰¹ See for instance (Sacks 1986) or (Ramachandran 2006) for a brilliant display of such cases.

5 CONSEQUENCES OF MY VIEW

The view I presented in Chapter 3 did not allow numerical persistence to account for our survival. If we are to accept this account of persons and lives, we should take a look at some of the concepts that are usually assumed to hinge on personal identity.

What follows is not meant to be an exhaustive account of all relevant concepts, but rather a cursory overview of some relevant concepts accompanied by rough suggestions on how we might deal with them. For all of these concepts we have the choice of either revising when they apply, or how we define them (i.e. tie them up to something other than personal identity). Taking desert as an example, we either conclude that desert does not require personal identity (you can deserve punishment and reward for things not strictly done by you) or we conclude that nobody ever deserves punishment or reward. The path of conceptual revision will be discussed in 5.1, while in 5.2 the question of whether we can manage without the given concept will be discussed – the question here becomes whether we need the concept for our current practices, and, if we do, whether we can be satisfied with alternative practices.

5.1 Conceptual Revision

Here follows a set of solutions following the strategy of redefining the relevant concepts in order to save them (in the face of simple persistence having been taken off the table)

5.1.1 Property

Property, we might argue, does not belong to the person here and now, but to the entire life. This matches poorly however with our idea that a person at any time is entitled to do whatever he wants with all his property. He can at any time give it all away and so take it out of the hands of his future. Perhaps we could patch this up by thinking of the current person₁ as a *representative* of the life with a mandate to manage all the property in fact belonging to the complete life. This amounts to a working (I think) but quite strange account of property. In this case the practical solution of 5.2.1 seems highly preferable.

5.1.2 Commitment

When we promise to do something later, but we will not strictly be there at the time, what are we in fact doing? Here we might successfully draw on the idea of the person as a character. Given that we are normally highly interested in our reputation, even beyond our natural death¹⁰², I think it is quite plausible to think of promises as tying a condition to a later evaluation of your character, specifically the trait of being a reliable person. If I promise to do something tomorrow, then, I am really claiming that I am the sort of character who would be moved to action by the commitments of my psychological ancestors (a person₁ in my past), and that I now make such a commitment for which my character will rightly be judged with regards to that specific trait. If my descendant should fail without good cause, he will reveal himself not to be a reliable character, and this, if I am concerned for the future of my life, will be bad for me. A person with an ancestry of broken promises will rightly be suspected of being an unreliable character and will consequently not be able to make credible promises. This is comparable, then, to promising that someone else will do something, where the force of that promise amounts to what reasons one has to think that this someone will do what you promise as a result of you making the promise.¹⁰³

5.1.3 Desert

A person is thought to only be justly rewarded or punished for his or her own actions, but if personal identity over time is out of the picture, this idea needs further interpretation. The simplest fix is to say that you are responsible, and so can justly be held responsible, for any action performed by your psychological ancestor. In some cultures desert was passed down through generations so that someone deserved a good or a bad treatment depending on whether his father or grandfather did good or bad. Even today in our own culture, we seem to think that a group or a people can deserve compensation for something they suffered in a time where all the current members had yet to be born.¹⁰⁴ We seem then to have a broader idea of

¹⁰² I do not mean to say we are all very conscious of our reputation, only that most of us would consider it to be bad to have a bad reputation and even to get a bad reputation after our deaths.

¹⁰³ Take the example of arranged marriages: Whether the father's promise of his daughter's hand in marriage has any real force comes down to whether his daughter is the type of character that is moved by an arrangement her father makes for her, or else whether he himself is the type of character that will force her daughter if she is unwilling.

¹⁰⁴ For instance a group like the Jewish people can be considered to historically "have gone through much discrimination, persecution and torment" and so because of that they may be considered to be deserving of some form of compensation either from "the universe itself" or more specifically from the groups that mistreated them.

desert not tied to personal identity but rather to a history – be it that of a person or a people. What is essential on this account of desert, then, is that the deserving person be of the same story as the person who did what the desert is concerned with.

This is close to what Lock says about desert. He ties both desert and personal identity over time to a unity which he calls consciousness.¹⁰⁵ Lock’s use of “consciousness” has little to do with the phenomenal experience or qualia which we often use it to denote today. Everything that you *can* recollect is for Locke a part of your consciousness. If the prince were to lose his old memories but could suddenly recollect the life of the cobbler, he would be of the same consciousness with the past cobbler and thus responsible for his actions.¹⁰⁶ You are, for Lock, responsible for everything you can remember doing, as you by definition are the same person as the one whose actions you remember from the first person perspective.

Whether this actually satisfies our needs for justification of judgements of desert is not clear. I suspect that this will seem, to most of us, more like a description of when we will ourselves find it psychologically required to be compensated or psychologically acceptable to be punished, more than an actual justification of this compensation or punishment. Perhaps it will be satisfactory for some.

5.2 Practical Adaptation

Here are the solutions following the other route. These will leave our concepts more or less the way they are, and instead look at what consequences the absence of personal identity over time has for the justification of our practices.

5.2.1 Property

Property is readily inherited. When someone dies, if nothing else has been specified their property is passed on to their closest family, especially their children if they have any. The continual transference of property downstream within a single life is, we might argue, nothing else but just such a convention of default inheritance. If nothing else is specified by the current person, the person who inherits his psychology also inherits his luxury yacht and his toaster. On the other hand, if he decides to give it all to some other person or organisation, the

¹⁰⁵ *Essay*, Book II, Ch XXVI “Of Cause and Effect, and other Relations” (Locke and Woolhouse 1997, 292-314)

¹⁰⁶ *Essay*, Book II, Ch XXVI §15-16 (Locke and Woolhouse 1997, 306-307)

person whom inherits his psychology has no special entitlement to this property. This would then compare to a dying person disinheriting his children. Even though they would have been entitled to his property if nothing had been specified, even though they were the default inheritors, they would then have no special entitlement.¹⁰⁷

5.2.2 Commitment

If commitments are held to be dependent on strict personal identity over time, then no one is ever bound by promises made in the past. But people who think they are, but nonetheless do not follow through reveal themselves as bad characters. People who realize that such promises are impossible should never *pretend* to promise anything, lest they reveal themselves as liars – another bad character trait. People who made promises in good faith but *then* read this thesis, however, must on this view be excused their promises. 5.1.2 might be the better choice here.

5.2.3 Desert

If we hold fixed that matters of desert hinges on personal identity over time, we must now conclude that desert is never applicable after all. The question then becomes: How well can we manage without it? I will split this question in two and talk first about punishment and the criminal justice system and secondly about rewards and virtues. At the end I will also discuss an alternative route.

Crime and Punishment

If a person never deserves punishment for any past action, what consequences does this have for our justification of punishment? We arguably need a legal system, but could we have rule-enforcement without appealing to desert? I think it will be useful to divide this discussion into something like calculated profit-driven crime on the one hand, and act-driven criminal behaviour on the other.¹⁰⁸ The idea is that while we still would need the disincentive of punishment to discourage a lot of the profit-driven criminal offences, in the case of crimes

¹⁰⁷ In Norway there are laws to restrict the disinheritance of your children (Arveloven: Lov om arv m.m. 1972 nr. 5 §29). Still something like this could be achieved if the property is given away before death. (John Asland, Note 184, lovkommentar, Rettsdata)

¹⁰⁸ This is not supposed to capture perfect, mutually exclusive, natural categories of crime. The idea is just that the crimes that clearly fit into either of these categories will need a different approaches; or so I will argue.

done impulsively or with just the act (not the result of the act) in mind, we can focus on the offender as having demonstrated a destructive or dangerous character. Let us take the calculated profit-driven crimes first.

Profit-Driven Criminal Behaviour and Psychologically Relatable Crimes

The sort of criminal behaviour I am imagining here is of a sort that society at large and so the judge and jury can psychologically relate to.¹⁰⁹ Tax-evasion, insurance-fraud, embezzlement, shoplifting, substance abuse, smuggling, copyright infringement, and similar behaviour in which the victim is either non-apparent or (at least seemingly) the offender himself. These are crimes where you typically have to follow the effects some steps along to find the real harm done. The laws against such behaviour include the rules that were typically first explained to you by your parents through an appeal to “Imagine if everyone did that” or “It’s the law, and breaking it is bad”. The selection mentioned is only meant as a rough guide though, and it might not be sufficiently curated – the point I am trying to make really only hinges on there being some behaviour we want to restrict without appealing to character flaws.¹¹⁰

The justification of punishing someone for some profit-driven crime would on this view be purely instrumental, as you would be the one punished not because you were the one that did the crime but because you are someone who profited from the crime, or more specifically: someone the crime was aimed at benefitting. Profit-driven crime is done for the purpose of benefitting your future, not your current, self. Knowing in advance that this future person is likely to end up with a fine instead of a bonus it starts to look like a pretty lousy gift. This can be compared to a dying person stealing money and giving it to his children. As he is dying he will not be deterred by the risk of punishment to himself, but he might be deterred if his children are the ones who will be punished.

Risk-takers might still gamble on not getting caught though, even if the instrumental punishments are set at devastating levels. The more important factor, then, will arguably be the probability of getting caught (as well as public awareness of this high probability).¹¹¹ If the probability of getting caught is close to 1, the punishment would arguably only need to

¹⁰⁹ Notice though that whether the crime is relatable through a story of how the character became the way he is is completely irrelevant here. The action itself needs to be relatable.

¹¹⁰ That is, this behaviour might to some extent imply a flaw in character, but it is a behaviour that we do not require that people abstain from for other reasons than the fear of getting caught.

¹¹¹ This is, of course, in the end an empirical question.. There is, however, research in this area that supports this claim (Mathiesen 1995, 89).

marginally outweigh the potential benefits of the crime. The next section will deal with behaviour we plausibly cannot prevent this way (and which we perhaps would not even be satisfied with having prevented this way).

Act-Driven Criminal Behaviour and Monstrous Crimes

If someone has raped, tortured or killed without there being some extreme excusing circumstance, we can reasonably conclude that the behaviour is indicative of a dangerous character. This is not behaviour the jury can identify with or relate to, but behaviour that really seems monstrous to them. Society (through laws, judges and jurors) could decide to lock up characters that have revealed themselves as capable of, or indeed even drawn to, such behaviour. The justification here lies solely in the prevention of further offences, and the goal must be the eventual rehabilitation and reintroduction of the character to society. If the character proves easy to rehabilitate in treatment, he can be released quickly. If he proves difficult to change, he might be kept indefinitely or eventually perhaps even “put down”. Notice that it does not matter if the jury can understand the behaviour of the offender in the context of his life’s history – that they can understand how someone can become so distorted in light of the terrible childhood they might have had. Equally it would not matter if they saw that this behaviour was understandable in light of a neurological disorder, say, or a brain tumour. All that matters is that the offender acted in a way that a well-functioning character would not, and that this scenario is liable to repeat itself. If the cause of the distorted character is a brain tumour that just means that the rehabilitation will consist in brain surgery rather than psychotherapy.¹¹² Before the offender is rehabilitated he must nonetheless be restricted from accessing the general population unsupervised. We could also imagine a person that when sober is a completely likable character, but that has a severe addiction to a substance that makes him a dangerous character. If he is unable to stay away from the drug that makes him a dangerous character, and the drug is readily accessible to him – that makes him a dangerous character when sober as well. This is then what he needs rehabilitation for – quitting the drug.

Staying with this category of criminal behaviour, let us now incorporate the thought experiments of fission (Appendix Chapter 4) to probe our conclusions further. If one of the qualitatively identical post-fission persons goes on to rape, torture, or kill someone shortly

¹¹² An astonishing example of a brain-tumour inducing bad behaviour is the well known case from 2002 where a man becomes a molesting paedophile from a growing tumour disturbing his orbitofrontal cortex. (Burns and Swerdlow 2003)

after the fission, we would on this view have good cause to lock them both up for prevention and treatment.¹¹³ The action itself is not what is important here, what is important is what the action tells us of the propensity for such behaviour in this character. If it was a long time after the fission occurred, and there were some relevant differences in the lives of the two post fission branches, say one of them had been traumatized by a long deployment in a warzone while the other one spent his days surfing, there would be less ground to arrest the surfer.¹¹⁴

In the end this line of reasoning means that if we had a perfect method of establishing someone's character and thus the propensity of this character to do something horrible under relatively normal circumstances, society could decide a limit of risk beyond which they would have as good a reason as they could ever have to lock someone up for treatment. The problem with this is just a practical one of whether we could ever make such a perfect test. As long as we cannot, the only good indication we have are the actions themselves – meaning we would have to await the first offence.

The justification of Norway's current penal code is to be found in the official document *Odelstingsproposisjon nr. 90*.¹¹⁵ It clearly states *prevention*, both of further offences and of first offences, as its main target and justification. In preventing first offences, the sentencing should work as a deterrent. In preventing further offences, the sentencing should work in a rehabilitating and/or debilitating fashion. Only through the need to appease society's need for justice (in order to maintain peace and order and especially to avoid vigilantism) is retribution drawn into the justification. This suggests that we are, at least formally or legally, not so far away from the view I am suggesting above. According to Robert Perkinson, the author of *Texas Tough: The Rise of a Prison Empire* the American penal system used to be anchored in rehabilitation, but that it has moved much more towards a retributive justification over the last 40 years.¹¹⁶

Virtue and Reward

The notion of someone having the right to be rewarded for some good deed is not as problematic. We of course want to incentivise some behaviour, and this we can do

¹¹³ I am assuming that the fission is perfect. In practice, there would of course be reason to investigate possible errors in the duplication-process as well.

¹¹⁴ See Appendix 4.5

¹¹⁵ (Odelstingsproposisjon nr. 90, seksjon 6: Hva er straffens formål og virkninger? 2003-2004, 77-81)

¹¹⁶ (Perkinson 2010)

instrumentally, just like with the punishment of for-profit crimes. Largely though, good deeds are done for their own sake, or for the sake of feeling good about oneself, or having people like you better, and this will all happen automatically. We do not have to drag rights into this. In fact, a lot of what can be seen as altruistic behaviour will seize if material compensations are systematically introduced – they get in the way of the original reward of a good conscience, and they seldom outdo it.¹¹⁷ If you care about your future self, you care about exhibiting a good character now, because that is what people will use to evaluate him, and rightfully so.

Responsibility and Past Actions

Finally there is something to be said about the notion of “taking responsibility for one’s actions”. While a person on my view never is strictly responsible for past actions, there is still a sense in which he should take responsibility. The person can still be thought to have a responsibility for shaping his future character in a positive direction. He should learn about the propensities and dispositions of his character from these past actions, and from this he should try to enforce the good traits and correct the bad traits. This in turn is what *regret* and *pride* can be interpreted as: the emotional basis for wanting to correct or enforce character traits revealed by past actions or behaviour. The person is justified in seeing the past action as indicative of his own character because the action was done by someone he has inherited his entire psychology from. He is right to feel bad, then, about the things his predecessor did wrong, and good about the things his predecessor did right, because he is right to assume that he has these positive and negative propensities too – and this he should respond to.

Because of this mechanism we can also think that we are to an extent licensed to *make* a person regret his predecessor’s bad actions – as well as rejoice in his good. This, then, gives an alternative justification for punishment and reward. You punish someone to make him regret and you reward him to make him proud. Like the approaches I discussed in 5.2.3, this too is purely instrumental and future directed, and so its justification hinges purely on its effect on the person in question, and on society. One reason for concern is that we might induce a regret of getting caught rather than a regret of the behaviour itself. (A regret of getting caught would of course be of no use.) Additionally, we would rather that the person

¹¹⁷ This effect is suggested by “Motivation Crowding Theory” discussed in (Frey and Oberholzer-Gee 1997) and (Frey and Jegen 2001). For a brilliant talk framing this topic see (Schwartz 2009)

regretted the behaviour on his own. Policing people's remorse might very well lead to an externalisation of this sentiment, which again might plausibly lead to a general weakening of its strength. Again, these are merely reasonable (to my mind) speculations on what must ultimately be settled empirically.

6 OBJECTIONS

These are some of the objections I imagine might crop up against my view.¹¹⁸

6.1 My Premises Are False (or Implausible)

It might be claimed that any of my starting premises are bad. I have not given a very good defence of physicalism, for instance, and I might certainly imagine objections directed against it. It is important to remember, however, that my argument only requires a very limited version of physicalism. (see Chapter 1.3.1)

It might further be objected that our folk-psychological view of what we are is a poor place to start the investigation. An animalist (Chapter 4.3), for instance, might argue that we should start with a biological conception of ourselves: that this clearly better defines what kind of beings we are. (That is, after all, what biology is all about!) The defence of my choice here can be found in chapter 1.3.2 and 2.3. It has to do with starting the investigation from a perspective that holds *value* to us, so that we might end up with answers that matter to us. (More on this in 6.3)

6.2 My Conclusions Are Unacceptable

This is the objection I can most easily imagine. Some of my conclusions are clearly *prima facie* counter-intuitive. It is a surprising claim, for instance, that we might survive through a future that we have no influence on.

6.2.1 Modus Tollens¹¹⁹

While my premises and reasoning *seem* perfectly acceptable, my argument is obviously unsound, it might be argued, as the conclusions I reach are completely unacceptable. Another way of putting this is that while there is some push towards accepting my premises, there is a greater push still towards dismissing my conclusions. If the argument appears to be valid then it could be interpreted as case of modus tollens: the valid argument shows, by reaching a false

¹¹⁸ It might be objected that what follows makes for a rather arbitrary assortment of objections, but then this is precisely the case.

¹¹⁹ $A \rightarrow B \ \& \ \neg B \vdash \neg A$

conclusion, that some of its premises must be false. It might for instance be argued that in the interest of preserving the coherence of some larger theoretical framework it might be worse to accept my conclusions than to deny some of my premises. Since I do not have a complete framework on offer, there is little I can do to meet this sort of objection.

6.2.2 Evil Consequences

It might even be argued that my conclusions and reformative suggestions are ethically bad, so that whatever analytical reasons we might have for accepting it is rendered moot. Conceptual neatness, it might be argued, can never compete with what is ethically good. To this I could say two things. First I could reply that truth is valuable in itself, regardless of its consequences. Secondly, I do not think that the suggested consequences of this view are ethically bad – I actually think they are good! I am again aware, however, that I do not offer a broader theory of ethics to back this up.

6.3 A Deeper Truth to Be Found

Perhaps some will think that my take on metaphysics means that while my thesis skates around on a conventionalist surface, there is a non-conventional and deep answer to be found underneath. As I have discussed earlier¹²⁰ there seems to be a conviction among a lot of metaphysicians that I have trouble wrapping my head around. The idea is that metaphysics can study the world at some deeper level, independent of our world view, deeper perhaps than even science can peer. I am completely dumbfounded by this position. While I agree that we can investigate our world view and adapt it to better fit scientific data or to get a more consistent picture, I cannot see how we could possibly hope to go beyond that.

Theoretical physics is in part a highly speculative enterprise. That is, it does a whole lot of speculating and theoretical model constructing. As with all proper science, however, the goal is to produce models that make solid predictions, predictions that can then be empirically tested. Without this restraint it would not be science.¹²¹ Metaphysicians can of course participate in this activity as well, but then this would be participating in the theory-building

¹²⁰ Chapters 1.3.1, 2.3, 4.2.3, and 4.3.4.

¹²¹ A popular objection to string theory, for instance, is that it makes no testable predictions but rather is able to accommodate any empirical data that could be found. (If this objection is true, string theory might perhaps merely be metaphysics?) See for instance (Woit 2006).

part of science; not something *beyond* science. Where I see metaphysics as potentially filling an unoccupied role is in the investigation and evaluation of our more basic ways of thinking about the world. When Sider says that the world has revealed itself as containing two different entities: particulars and properties, I am not sure how to properly interpret him.¹²² What I want him to say is that we have revealed something about the way we think about the world: we divide the world into particulars and properties. From this I can immediately see both a descriptive inquiry on what these categories imply and how they relate, as well as a normative inquiry into whether these categories can be replaced by something better (whether we can find a better way of slicing the pie). My concern is, however, that this is not what he means.

6.4 Are Persons₁ Temporally Extended or Not?

My account appeals to persons at times, or in situations, as persons₁. I claim that this conception of a person is one under which the person should not be thought to persist. At the same time I claim that a person₁ cannot be thought to exist in a literal *point in time*. The person₁ is supposed to experience, to hold a perspective, to think, and to act. This means that a person₁ must have a temporal extension – but then why not for an entire lifetime?

This is actually one of the points of my account that I feel needs more work. I think (and hope) that the relationship is sufficiently clear for my account to work, but that does not mean that I think it is explicitly clear at all. I am unsure of precisely how I should account for the relation between the conception of a person as subject and the conception of a person as a physical object. What I suspect is going on is that these ways of grasping the world are not easily translatable. My main claim is perhaps that we cannot re-identify persons₁. I have included a thought-experiment (A5.11) in the appendix to probe this question of subject-duration. When the frequency is very high, there is arguably either no subject present at M2, or the same subject is present there as is present at M1. There is a certain pull, especially in version two of the experiment (where the updates from M1 are executed as adjustments rather than replacements), towards saying that there might be experience going on at M2 as well. But there is certainly not a distinct functional subject at M2 – this “man” does not find himself in the room he is in, he cannot perceive anything of this situation, and he clearly has no perspective from which he can act in this room. As the frequency is lowered, the individual updates gradually has more and more time to respond to the room. When the updates get far

¹²² I quote a part of this in chapter 4.3.4. (Sider, Hawthorne, and Zimmerman 2008, 1-3)

enough apart there is clearly a distinct subject present, a different person₁ than the one in the other room.

I think a plausible analysis is this: “Person₁” picks out the conception we use when orienting ourselves in a situation. It is the concept we use for picking out who this situation concerns. “This is what *I* have to deal with.” It is also used for thinking about others in the same way: “This is what she had to deal with.” It is this subjective perspective that I want to pick out, of being in a situation, receiving certain impressions, deliberating, making decisions, and acting. (And not necessarily all of these elements for any given situation.) *How* we react and deal with a given situation has to do with what we are *like*. If we remove this from the equation (create multiple persons₁ that are identical persons₂), we are different only in holding different points of view; but these points of view are all equal. It is no fact about the world that I am me and you are you. It is the same perspective that we all employ, for ourselves, to orient ourselves to the situations that we through this perspective find ourselves in. Something like that. Now, how long does a person₁ last? I don’t know.

6.5 There Are Better Accounts

It might be objected that my account does a poorer job at what I set out to do than some other account X. This is of course not only possible, but even very likely. There are countless accounts out there, and I am in no way acquainted with them all. Neither am I confident that my account, as it is presented here, perfectly captures my view, or that it gets all the details right. The account I have presented is not defined relative to any specific system of ethics, epistemology, truth, or language. Other accounts may be more firmly grounded. I do believe, however, that my main view stands a fair chance of being assimilated into such a larger system of thought.

7 CONCLUSION

7.1 What I Have Tried to Accomplish

Ultimately, it is up to each and every one of us to identify ourselves as the kind of thing we think we are, and also to identify what about this existence we care about. What I have tried to accomplish with this thesis is not so much an argument for the one true way of thinking about our lives, rather I have tried to present a consistent view that is close to what I take to be the central intuitive view we have of ourselves, at least in the culture with which I am familiar.

In broad strokes, my main argument has been the following: While the intuitive view might be fine for the practical purposes of *living* a life, it makes for a poor conception of what our survival can be thought to consist in. In going after a better account we have to be clear on what it is we are after. What is needed is an account that relates the basic folk psychological truths about ourselves (as being agents and experiencers) to our scientific understanding. This approach is needed in order to produce an account that can have any bearing on our lives. On the conception we readily have of a person when we consider him at a time, the person is a numerically distinct subject,¹²³ but this numerically distinct entity cannot be thought to be what persists over time to form a life. However, the considered person can be thought to have a history and future, and in this way be *engaged* in a life – *or in several*. While we might be tempted to have “person” designate whatever might coherently be said to persist through such a life, and focus our efforts in accounting for these persistence conditions – that (if achievable) would mask this very real way in which the person does *not* persist. The person construed as an agent has a distinct reach into the future if there are one or more persons in the future that inherit his psychology. This relation allows for a future oriented life-project to be maintained, as well as a retrospective life-narrative.¹²⁴ While the life-narrative can be thought to contain the person’s psychological identity, or self-image, the life-project can be

¹²³ Under this conception the person is both an agent: someone who acts, and an experiencer: someone who experiences. He is an acting and experiencing subject.

¹²⁴ Marya Schechtman (Schechtman 1996) advances a view of personal identity where the self-told narrative is the basic unit. It would be interesting to see whether her account might be compatible with mine, as I understand her to account for how our lives are woven together, rather than what underlying relations makes this possible. Whether a merging of our accounts would be beneficial (or even whether it would make sense) is still unclear to me.

thought to contain much of the value for the person in having a future.¹²⁵ This allows even a person that does not have any reach into the future to survive by having someone else maintain his life-project. This is the case for the thought experiment of the Soldier/Surfer (A4.5), where the post-fission person that shoots himself can survive through the surfing-future. While he has no reach into this future, the life-project he is engaged in will be maintained by it. The person construed as an *experiencer* on the other hand has no extension into the future. Neither the surfer nor the soldier will in any deep sense get to experience this future surfing. Anticipation in general still has its function preserved though, as the future you anticipate is the future your psychological descendant will be prepared for. The Soldier cannot anticipate the surfing future on these grounds, but he can still “look forward to it”.¹²⁶

7.2 Further Impact

In addition to the suggested practical consequences of 5.2, there are several aspects of our lives that I think might potentially be impacted by embracing the view I advocate in this thesis. How we think about the grand sum of our existence, about our death, and about the promises of technological routes to immortality.

7.2.1 “Future People Equality”

There is a distinction in my account, between the construal of an agent and that of an experiencer. Our agency stretches into the future, but our experience does not. When we think of our future, we should think of what we will do. When we prepare for our future, we should prepare *as if* we are to experience it. But when we consider the value of what good or bad things will be experienced in the future, we should think equally of everyone who will be there. When we consider the value of our own future relative to that of others, we should think of what role we will play. We will get to experience no more and no less of the future regardless of what happens. It should be important for us to have a psychological continuer only insofar as there is something in that future that we wish to do, or a role we wish to fill, and then only insofar as there is no one else we can trust to do this action or fill this role.¹²⁷

¹²⁵ Though many of our plans and intentions that also give our future value are much more local, and not tied up to this kind of long-term life-project.

¹²⁶ I here use “look forward to” as discussed at the end of 3.1.2.

¹²⁷ Again, this is not to say that there is no intrinsic value in there being such a future life. The point is that the special value that this particular future life has to you, is relative to your roles and intended actions.

Wishing to be there for your child, filling the irreplaceable role as a mother, say, makes all the sense in the world on this view. Wishing to literally just “see your child grow up” on the other hand, does not.

7.2.2 Death

I share Parfit’s sentiment on how a reductive perspective changes one’s outlook on life, and so on death. I think much of people’s positioning in matters of religion as well as metaphysics can only be understood from the perspective of mortality. There seems to be something almost incomprehensible about one’s own death, and yet we all know we will die. Well, some of us think they will not die – they think they will live on for all eternity in some form or another. It is easy to see how such a view might be motivated by the perceived impossibility of imagining one’s own death – or perhaps the unbearability of it. But so we can understand the motivation for the position I defend as well. I get out of the bind, not by inflating an afterlife, but by deflating my current life. Life, on my view, is not so impossibly different from death – not because death is so much like what we took life to be, but because life is so much like what we took death to be. Our simple hope of getting to experience the future never comes true – just as we imagined this to be a hallmark of our dreaded death, it has instead become a truth about our lives. We never get to see the future in this naïve sense, but that’s okay. It’s not like we lost something – it was never there to begin with.¹²⁸ Now that we realize this, we can get used to the idea of dying – not because there is more for us waiting on the other side, but because there is not any *less*. I am not arguing that death is not a bad thing – that would be foolish. There are many things bad about our death – there just isn’t this extra unfathomable badness on top. Death does not mark the end of a lifelong window of experience. Death is not the end of a subject of experience. Say it out loud.

7.2.3 The Singularity

As all our micro-transistor related technology is getting exponentially smarter over time, and more and more of our technology is becoming micro-transistor based, we are getting used to ever increasing rates of improvement and obsolescence. Predicting the state of technology ten or

¹²⁸ I do agree with Nagel, that this “Parfitian” survival is in a sense worse than what a non-reductionist survival would offer (Nagel 1989, 224). I do not, however, agree that this picture of survival is depressing. Rather, I agree with Parfit when he describes this insight as liberating (Parfit 1984, 281). I guess this difference of opinion is for psychologists to work out.

twenty years ahead has almost become an absurdity. Still, some people eagerly spend their time scrying all the same, and some of these futurists are convinced that artificial intelligence will surpass human intelligence within decades. Not because it is any good now, but because it is getting better at an exponential rate.¹²⁹ I am in no position to evaluate the claims of the futurists. They seem improbable, but then I am reminded of what our digital technology was like twenty years ago. Speculations on potential benefits like achieving immortality through the “uploading of our consciousness to machines” do however give us yet another reason to think long and hard about what it means for us to survive (and what about it that makes it so desirable). If these futurists are correct, this thesis might see more practical application than is perhaps usually granted a metaphysical inquiry. Perhaps, though, this is because this thesis is not properly a metaphysical inquiry after all. It is, however, the deepest sort of inquiry I can conceive of.

¹²⁹ See (Kurzweil 2006)

Appendix : Enumerated Thought Experiments

Herein follows a list of thought experiments meant to demonstrate the difficulties we face in accounting for the nature of our survival. The thought experiments are mostly presented without any claims to how we should evaluate them, leaving the conclusions for the reader to make. For my own analysis of these scenarios, turn to chapter 3 (mainly 3.6) in the main part of the thesis. Also, for the scenarios I am aware of stealing wholesale from other writers, I will credit this in footnotes to the respective headings. For the rest I either made them up, or have been unable to track their origins. I apologise in advance for any unaccredited thefts I might be committing.

The list is sorted in five categories based on what kind of challenge they hold for our account. There is also a general tendency of increasing outlandishness throughout the list; both of the categories and of the scenarios within each category. We start out easy.

A1 Surviving Accidents

A forty year old woman is involved in a car accident, she sustains severe damage from the collision and when she wakes up in the hospital several days later:

A1.1 Bodily Damage

She has lost both her legs and will have to get used to riding a wheelchair from now on. Her family is sorry that she cannot walk anymore, but they are glad she survived and thankful she escaped without any brain damage.

A1.2 Lost Memory

Her body seems fine, but she has lost every episodic memory she had. She cannot remember ever having experienced anything before, and additionally she cannot recognise any of her family members. She seems to the family members to be pretty much the same familiar person they knew so well, but it pains them to see that she now regards them as strangers in return. She will try to get to know them again, and they are comforted in knowing that she once loved them, and that they already know the ins and outs of her personality. They are convinced that she will come to love them once more.

A1.3 Altered Personality

Her body seems fine, and she can remember her life just fine, but she no longer identifies with the person in her memories. She no longer wants to live the life she has been living up to that point. Her family no longer recognise her in the way she acts, and feel as though she has become a different person altogether. They find comfort in thinking that although the woman they all knew and loved will not return to them as she was; she survived to live a different life — a much better fate they conclude, than had she died. Still, the period of getting used to this new personality will be painful for everyone.¹³⁰

A2 Suspended Life

A2.1 Hypothermic Conservation

After realising that it when it came to patients that had been very cold, it was possible to resuscitate them even when they had been completely lifeless for long periods of time, the rule in Emergency Rooms became "No one is dead until warm and dead".¹³¹ As of 2010 this effect is actively used to allow surgeons to perform operations that require the blood flow of the patient to be stopped for the duration of the operation.¹³² The phenomenon is accurately described as "suspended animation": All life functions end – there is no detectable brain activity, the heart is not pumping, and the lungs are not breathing. The fact that you can survive several hours of being in a state like this precludes a conception of death as the seizing of all vital functions, and pushes instead for a conception of death as having sustained irreparable damage. What medical conditions will be synonymous with death will under this conception clearly be dependent on the state of the medical science at the time.

A2.2 Cryogenics

It is no big leap of imagination to see how this method of hypothermic conservation could be further advanced, and ultimately to a point where the state of a body in principle could be preserved forever. A person with a terminal illness could then deep freeze himself and be thawed successfully back to life in some distant future where a remedy has been discovered.

¹³⁰ This sort of damage is what is suffered by the victims of frontotemporal dementia. Unlike Alzheimer's disease the patients lose not their memories but rather their personality and their self monitoring capability: "FTD alters the personality and patient's self [...] and political, social or religious values. Yet FTD patients lose their awareness for these changes due to deficits in self-monitoring" (Schroeter et al. 2007)

¹³¹ (Iyer et al. 2007)

¹³² (Back From the Dead 2010, BBC HD)

A2.3 *Clumsy God I*

Kim is having guests over for dinner. He is standing in the middle of the room, engaged in conversation with several of the guests, when he suddenly explodes into a cloud of particles. The talking comes to an abrupt halt as everyone is now staring with disbelief at the space that just moments ago contained their host for the evening. They are not left standing this way for long however, as a strange figure of a man soon appears in the room next to the barely visible mist that was once Kim. The strange man immediately starts apologizing, saying something about taking a wrong turn somewhere, but importantly he promises to rectify the damage he has caused. He turns to the fog of particles that is now evenly dispersed over large parts of the room. He seems to be concentrating. Soon they can all notice the fog starting to gather around the spot where Kim once stood; at this point the strange man gives a brief nod to the guests and disappears just as abruptly as he appeared. The cloud of particles now visibly takes on the shape of a man and before long Kim is once again standing in the middle of the room. He starts completing the sentence that the sudden event had interrupted him from completing, but is quickly stumped by the looks on the faces of all his guests. "What? What's going on? Why are you all looking at me like that?"

A3 Replacing Parts

A3.1 *A Baby Grows Into an Old Man*

A three year old baby grows into an eighty year old man. All his cells and all the molecules constituting his cells have been replaced many times over. All his memories from the time he was three are gone. His desires, his preferences and much of his personality have changed many times over.

A3.2 *The Island*¹³³

Sometime in the not so distant future, some very rich people have decided to keep clones of themselves locked away as a safety measure, in case of any injury. If one of them loses a limb or damages an organ beyond repair, a spare can quickly be acquired from his or her clone. Some of these rich people have even started swapping out parts of their body with parts from younger clones in an effort to stay young. In principle any part can be replaced, and some of the more extreme users have swapped almost every part of their bodies at some point.

¹³³ This is pretty much the plot in the movie "The Island" (www.imdb.com/title/tt0399201) directed by the always terrible Michael Bay.

A3.3 Brain Transplant

One of the clone-insured people in the above scenario decided to replace his brain, as he wanted a younger and more agile mind and wasn't too fond of his old memories anyway.

A3.4 Joker God

A woman sits in her office chair, typing out an email on her computer. A mischievous god has been keeping an eye on her for a while. The god decides to suddenly replace every single atom constituting her body with qualitatively identical atoms. This happens so fast, that the man sitting across from her, although he is looking at her when it happens, registers no change in the woman. She continues to type out the email as if nothing had happened. Later that day they will be going out for coffee; the man has not completely forgiven her for standing him up the last time.

A3.5 Clumsy God II

Like “Clumsy God I” (A2.3), but in this scenario the strange visitor does not use the particle cloud to recompose Kim, rather he fashions him out of particles extracted from various places around the house. The cloud he simply waves away.

A3.6 Teleportation¹³⁴

A traveller enters a machine on earth and presses a green button. His atomic composition is instantly scanned after which he is pulverized, his atoms stored for later use. The information of his composition is sent by radio waves to another one of these machines stationed on Mars. There the information is used to compile, with matter stored locally (most of which is from travellers who at some point left Mars), an exact duplicate, atom by atom, of the traveller who entered the machine on earth. The traveller emerging in the machine on Mars will experience his journey as having been instantaneous; of course he knows that some time must have passed to allow the information to reach Mars, but he feels like his surroundings have just been instantaneously swapped out.

¹³⁴ Or “teletransportation” as Parfit calls it. This is one of Parfit’s famous thought experiments, or at least close to it. For his own version see (Parfit 1984, 199). The idea is also contained in pretty much every episode of the TV-series “Star Trek” starting from 1966.

A4 Fission

A4.1 Hemisphere Split¹³⁵

Three brothers are involved in a car accident. They are triplets. Two of them sustain irreparable brain damage whilst the third crushes parts of his spinal cord. The ER-crew manages to preserve the bodies of the two brain-dead brothers and the brain of the brother with a broken body. For some reason (obviously for pimping out their CVs) the surgeons decide to transplant one hemisphere of the healthy brain into each of the two excavated skulls of his brothers. Several weeks later the two resulting patients wake up, each thinking he is the brother whose brain survived, but seeing themselves in the mirror they are shocked to see that they appear to be trapped inside the body of one of their brothers. Upon seeing each other though, and then learning what has actually happened, they are not sure what to make of the situation at all...

A4.2 Teleportation with Several Copies

Building on the teleportation scenario described in 3.6: This time when our traveller presses the green button, the information about his composition is sent not to one place but to two. The two emerging travellers have no idea at first that they are not the only copy around, but one day they bump into each other on the street.

A4.3 Teleportation with Delayed Execution

Upon pressing the green button our traveller is surprised to find that he is still standing in the machine he entered. A technician is quick to inform him over the loudspeakers that they have been experiencing some problems with the teleportation unit he is in, but assures our traveller that his copy has already been materialized on the other end (he is just going across town this time), and that if he'd be so kind as to just wait a while, he promises to send someone down there to take care of him...

A4.4 ECC Teleportation

Following a media storm over some "unfortunate incidents" (as the Tele-Trans Company calls them) a new version of the teleporter-machine is launched. It utilizes an error checking and correction routine that compares the finished copy with the original before wiping the original. This means that there will be a delay, its duration dependent on the distance of

¹³⁵ This is a version of the second famous thought experiment from Parfit. For his version see (Parfit 1971), he places its origin with Wiggins, who again cites Shoemaker.

travel, in which the traveller sits in the machine and waits for confirmation of his successful arrival at the selected destination, upon which he is disintegrated. Some of the local branches of the Tele-Trans Company even thought it amusing to put in a camera so that travellers on the departure side could see themselves emerging on the receiving end. This practice was soon discontinued however, as a lot of travellers were disturbed by the sight of themselves from the outside – some reported experiencing the evident failure of the machine to transport their consciousness along with their body, as they were "obviously still *here*". This did not cause as big a spectacle as one might think though, because the traveller who arrived at the recipient end of the teleportation had, of course, no memory of feeling this way, as this would have transpired after the scan had been made.

A4.5 *The Soldier/Surfer*

You are told there will be made two copies of you in a year's time. One of these copies will be sent behind enemy lines as a specialist soldier; the other will be given a surfboard and a three year grant for surfing full time. These copies will be made using you, at that time, as the template; the template itself will be destroyed in the process. One copy will wake up in army gear and rifle in hand; the other will wake up holding a surfboard. You will have to decide how to deal with this situation, and what you should prepare for.

Can you hope to become the surfer and not the soldier? If not, should you prepare for both futures equally? What about simply skipping the military training and just prepare to suicide if you find yourself holding a rifle after the fission? What makes that a bad option if you did not want the soldier-future in the first place? If you wake up with a rifle, but want to go surfing, what reason could you have for shooting the other copy and stealing his board, over just simply shooting yourself?

A4.6 *One Way Mission*

You are contacted by the government. They have been struggling for some time with a critical malfunction in one of the unmanned space stations in orbit around Uranus. The station functions as a control relay for an important experiment in progress at the planet surface. They have finally been able to identify the problem, but it requires sending someone out there to fix it; and it turns out that you are the only one qualified to make the repair. The trouble is that this will be a one way mission, as the required timeframe makes it impossible to arrange for a return journey. The government is however not about to ask you to abandon your life for this mission, rather they ask you only for your consent in sending a physical duplicate of

you. This would require you to report to their research facilities where you will be briefed on the mission and then scanned. After that you will be free to go back to your daily affairs. Because the duplicate will be assembled on the space station, you will have to be briefed on the mission before the scan is made.

A5 Beyond

A5.1 The Clinical Trial

You decide to participate in a clinical trial of a new drug that has proven to effectively erase the next ten minutes of the patient's life. This includes not just every memory created in that period, but also any skill learned as well as any change in preferences and habits. You agree to participate in the trial, thinking you can do without ten minutes of your life, and the reward for participating is substantial. Just as you are about to receive the drug you realize that as long as you go through with the injection, whatever great ideas or sudden realisations you have will be lost. You will be psychologically isolated from your own future. You start to get anxious and hope that it will not be now that you finally muster the willpower to stop smoking, or take up writing, or anything like that. You eventually calm yourself down by telling yourself that this really isn't much different from getting very drunk, and you decide to go ahead with the injection.

A5.2 Clumsy God III

Like A2.3 or A3.5, except that in this scenario the strange traveller recreates Kim, not as he was immediately prior to the incident, but rather exactly as he was just ten minutes before – Kim was outside talking to his mother on the phone at the time. As he appears before them he is completely stunned at how the entire scene seems to have shifted around him. His friends tell him what happened and it is soon clear that he has no memory at all of the last ten minutes. Other than that he seems fine.

A5.3 Awaiting the Announcement

You are standing beside your exact duplicate, awaiting the announcement. You have just been produced there, you do not know if you are the original template, or indeed if any of you are – the template may be somewhere else. What you do know is that the two of you standing there now are exact physical duplicates as of only seconds ago. You want to live, you want to go home and be with your loved ones, but you know that he is thinking the same thing, and about

the same people no less. Soon the announcement will ring in the speakers and only one of you will be let out, the other is destroyed. As you stand there now, awaiting the announcement, the question is what it always is in this game: Do you have any reason to prefer one outcome over the other?

A5.4 Whimsical Gods (Providential Survival)

A woman is sitting in her office chair, typing up emails. A man is sitting across from her, looking at her. A whimsical god has decided to remove her from existence: He wills away every atom in her body and she is gone. Completely unrelated to this, another, equally whimsical (though more creatively inclined) god turns his attention to the now empty chair and imagines a woman sitting in this chair typing up emails. He wills this woman into existence and it just so happens that the woman he created is qualitatively identical to the woman just willed away by the other god who was just there. All this has transpired so quickly that the man sitting across from her, although looking straight at her, could not detect any of it. To him it seems that the woman he knows so well is still sitting in her chair, typing up emails.

A5.5 In a Galaxy Far, Far Away...

In a galaxy far, far away, there is a world very much like ours, inhabited by exact duplicates of the people walking around here.¹³⁶ You are told that even if you die here on earth there might still be a duplicate of you living on in some distant part of the universe. Can this be of any comfort to you?

A5.6 ...A Long Time Ago

Perhaps a person existed a long time ago that was qualitatively identical to how you are right now, and in a qualitatively identical environment to the one you now live in. If this is so, and he lived on from that time to become an old man, would it be less bad for you to die now? The idea then, is not that your life will go on happening somewhere else, but rather that it in some ways already did.

A5.7 The Generic Sufferers

An extremist group of experimental philosophers has started capturing people and turning them into generic sufferers. They have perfected a procedure that reduces a person to a simple creature stripped of all the features normally thought to “make us who we are”. While the

¹³⁶ For an argument that this may indeed actually be the case see (Rucker 2005)

post-procedure creatures maintain basic behaviour like feeding and pain avoidance, their higher cognitive skills along with all their memories are lost.¹³⁷ These creatures are locked away from the public in hidden facilities where they spend their days strapped to machines specifically designed to torture them without end.

You have just been captured by this group. You will soon be processed, and what will result is yet another one of these generic suffering creatures. The question becomes: Do you have any reason to dread the future experiences of this particular creature, any more than the experiences of any of the other creatures. Do you have special reason to dread the experiences of all of these creatures, more than anyone that has not been targeted or captured by the group? Or should you simply consider this to be your death? Will your friends and family have special reason to pursue the termination of this particular creature's miserable existence? (These questions are of course all different ways of making vivid the question of whether this creature will be you in the relevant sense – whether this creature's suffering will be *yours to suffer*.)

A5.8 Generic Sufferers 2

In this scenario the suffering creature is made from you, but this time the process involves a stage where your entire body is decomposed to a heap of particles – much like in the teleportation scenarios. The generic sufferer is then built using the matter you once consisted of.¹³⁸

A5.9 Body Swapping Twins

In the not too distant future, a pair of monozygotic twins pilot a new research project. At an early age their entire neural network is mapped out with the goal of uncovering where all the nerves in the neck come from and where they go. Using this map they go in at some point in the neck and cut all the nerves so that they can rearrange them to a common configuration and reattach them via two connector boards. In addition to all the nerves, the arteries, vessels, digestive tract and wind pipe are also cut and reconnected in this way. The result is a “plug and play” joint in the neck of the two twins where they are identically structured, effectively rendering their heads (or their bodies) modular components. The point of all this is that the twins will, throughout their life, frequently be exchanging their heads (or bodies). As their

¹³⁷ Perhaps the procedure involves something like removing the majority of the cerebral cortex, leaving behind the parts of the brain we have in common with the reptiles.

¹³⁸ The degree of similarity in structure of this creature and the corresponding parts of yourself can be also be varied.

nerves are remapped correctly at the joint, their heads require very little time adjusting to the new body after an exchange. Once a year they report to a clinic where the heads of the two are exchanged. At the age of forty they are asked to account for their lives so far and for their plans for the future. What would this account look like? If they identified with their head, tracking its journey as their life, could they somehow be wrong in their account? If one of them commits a crime at some point, but the crime is first discovered after a swap has occurred, which one do we punish? If one of them is sentenced to several years of prison how can this time be served if the swapping is to continue?

A5.10 A Gradual Replacement

An internationally wanted criminal goes to extreme measures to hide from the law. He figures his only chance is to disappear for some fifty years, but he does not want to waste away hiding in a basement. He decides to be the guinea pig in the pilot run of his cousin Tony's great new invention. It is a machine that perfectly scans the physical make-up of a person right down to the atom and stores the atomic configuration as an image, much like the teleportation devices in the foregoing scenarios. Unlike the teleportation devices however, this machine does not send that image to another machine that immediately constructs a copy from raw materials. Tony's machine is much more subtle – and cruel. It is designed to store the image file for a long time, and then, when the time comes, to gradually impose that captured structure on another living person. As the nameless criminal (Frankie) enters his cousin's machine he is scanned and vaporized, his composition recorded and stored. Now the information just sits in a hard drive for fifty years, as was the agreement. After fifty years an unsuspecting man is conned into attending weekly "personal development" meetings at Tony's facilities. He is told he will be testing out a new machine that "improves your character and rids you of your shyness and insecurities". What in fact happens is that every time the man comes in for his weekly sessions and sits in the machine, his neurological structure is gradually forced into alignment with that of the recorded criminal. As the weeks pass he becomes less and less like the man that was conned into treatment and more and more like the criminal mastermind that entered his cousin Tony's machine all those years ago. In the end the man has no memories but those of the criminal, no character traits but his, no goals, drives or concerns of his are connected to the life of the victim. It appears that the victim is gone and the criminal has returned – Tony Carbone greets him as his nephew.

A5.11 The Frequency Modulator

Two neighbouring rooms in a future research facility each contain a platform. These two platforms are connected. They are named M1 and M2. The purpose of M1 is to scan whatever object is placed on it and send this information to M2. M2's purpose is to reproduce the object scanned by M1. This M1-M2 setup is able to execute this whole operation almost instantaneously, and it has the ability to operate continuously – constantly updating the object on M2 by what is scanned on M1. There are two versions of this experiment. In one version M2 always materializes its product from the ground up, disintegrating whatever is already on the platform. In the other version, M2 first creates the object from ground up (a key frame, as it is called in video encoding) but consequently only updates the object by *rearranging* it to correspond with the object on M1.

In the first room there are two people in conversation. One of them is standing on M1. The frequency of the M1-M2 updates are set to 100Hz (100 updates per second). You walk from the first room where the two people are talking and into the second room – the room holding M2. On the platform marked “M2” there stands what looks to be the same man from the first room, and he looks to be in the same conversation. It appears that he is looking at someone, the man he is in conversation with, but there is nobody in the room but him and you. You try to get his attention, but the man does not respond to you, he does not look at you, he appears more than anything to be a holographic video-feed of the person in the other room. Then, the frequency of operation is gradually lowered. You start to detect jitters in the motion of the man. And after a while, when the frequency has dropped to only a few updates per second, it starts to look as if he is constantly interrupted from completing a motion. When the updates get even further apart, you see him get a confused expression on his face, over and over again, always jolting back to an updated position. As there begins to be several seconds between each update, the man clearly acknowledges you, and you start to interact with him. “Hello”, you say. “Hi”, he replies, “How did I get..”, but he is interrupted again. This conversation repeats itself. Each time it gets a little longer. You start to see these updates as clearly distinct. You can no longer see the man in conversation on platform M1; all you see are distinct updates. Each time there is an update, your interaction begins anew.

Bibliography

- Back From the Dead. 2010. BBC HD, Original broadcast date: 27 September 2010, Television.
- Bohn, Einar. 2011. "The Logic of the Trinity." *Sophia* no. 50 (3):363-374. doi: 10.1007/s11841-011-0265-1.
- Burge, Tyler. 1986. "Individualism and Psychology." *The Philosophical Review* no. 95 (1):3-45.
- Burns, Jeffrey M., and Russell H. Swerdlow. 2003. "Right Orbitofrontal Tumor With Pedophilia Symptom and Constructional Apraxia Sign." *Arch Neurol* no. 60 (3):437-440. doi: 10.1001/archneur.60.3.437.
- Cartwright, Nancy. 1993. "In Defence of 'This Worldly' Causality: Comments on van Fraassen's Laws and Symmetry." *Philosophy and Phenomenological Research* no. 53 (2):423-429.
- Chalmers, D.J. 1996. *The conscious mind: in search of a fundamental theory*: Oxford University Press.
- Chen, I. 2010. "When Character Crumbles." *Scientific American mind* no. 21 (5):30.
- Frey, Bruno S., and Reto Jegen. 2001. "Motivation Crowding Theory." *Journal of Economic Surveys* no. 15 (5):589-611. doi: 10.1111/1467-6419.00150.
- Frey, Bruno S., and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out." *The American Economic Review* no. 87 (4):746-755.
- Gendler, Tamar Szabó. 2002. "Personal Identity and Thought-Experiments." *The Philosophical Quarterly* no. 52 (206):34-54. doi: 10.1111/1467-9213.00251.
- Gilbert, Mads, Rolf Busund, Arne Skagseth, Paul Åge Nilsen, and Jan P. Solbø. 2000. "Resuscitation from accidental hypothermia of 13·7°C with circulatory arrest." *The Lancet* no. 355 (9201):375-376. doi: 10.1016/s0140-6736(00)01021-7.
- Harlow, J. M. 1848. "Passage of an Iron Rod through the Head." *The Boston Medical and Surgical Journal* no. 39 (20):389-393. doi: doi:10.1056/NEJM184812130392001.
- Iyer, Anand, Vincent Rajkumar, Dinesh Sadasivan, Jaye Bruce, and Ian Gilfillan. 2007. "No one is dead until warm and dead." *The Journal of Thoracic and Cardiovascular Surgery* no. 134 (4):1042-1043. doi: 10.1016/j.jtcvs.2007.05.028.
- Jonsen, A.R., R.M. Veatch, and L.R. Walters. 1998. *Sourcebook in bioethics*: Georgetown University Press.
- Kim, J. 2006. *Philosophy of mind*: Westview Press.
- Kurzweil, R. 2006. *The singularity is near: when humans transcend biology*: Penguin.
- Lewis, D. 1976. "Survival and Identity." In *The Identities of Persons*, edited by A.O. Rorty. University of California Press.
- Lewis, David K. 1968. "Counterpart Theory and Quantified Modal Logic." *The Journal of Philosophy* no. 65 (5):113-126.

- List, C., and P. Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*: Oxford University Press.
- Locke, J., and R. Woolhouse. 2004. *An Essay Concerning Human Understanding*: Penguin Books Ltd.
- Locke, J., and R.S. Woolhouse. 1997. *An essay concerning human understanding*: Penguin Books.
- Mathiesen, T. 1995. *Kan fengsel forsvares?:* Pax Forlag.
- McLaughlin, Brian; Bennett, Karen. 2010. "Supervenience." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Mitton, Jeffry B., and Michael C. Grant. 1996. "Genetic Variation and the Natural History of Quaking Aspen." *BioScience* no. 46 (1):25-31.
- Moyer, Mark. 2008. "A survival guide to fission." *Philosophical Studies* no. 141 (3):299-322. doi: 10.1007/s11098-007-9161-5.
- Nagel, T. 1989. *The view from nowhere*: Oxford University Press.
- . 1991. "Death." In *Mortal questions*, edited by T. Nagel. Cambridge University Press. Original edition, *Nous*, IV, no. 1 (February, 1970).
- Odelstingsproposisjon nr. 90, seksjon 6: Hva er straffens formål og virkninger? 2003-2004. Tilråding fra Justis- og politidepartementet av 2. juli 2004, godkjent i statsråd samme dag. (Regjeringen Bondevik II).
- Olson, E.T. 1997. *The human animal: personal identity without psychology*: Oxford University Press.
- . 2003. "An Argument for Animalism." In *Personal identity*, edited by R. Martin and J. Barresi, 318-334. Blackwell.
- . 2008. "Reply to Zimmerman." *Abstracta* (Special Issue 1):38-42.
- Olson, Eric T. 2010. "Personal Identity." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Papineau, D. 2004. *Thinking about consciousness*: Clarendon Press.
- Parfit, D. 1971. "PERSONAL IDENTITY." *Philosophical Review* no. 80 (1):3-27. doi: 10.2307/2184309.
- . 1984. *Reasons and persons*: Clarendon Press.
- . 2007. Is personal identity what matters? *The Ammonius Foundation*, <http://ammonius.org/assets/pdfs/ammoniusfinal.pdf>.
- . 2008. "Persons, Bodies, and Human Beings." In *Contemporary debates in metaphysics*, edited by T. Sider, J.P. Hawthorne and D.W. Zimmerman, 177-208. Blackwell Pub.
- Perkinson, R. 2010. *Texas Tough: The Rise of America's Prison Empire*: Picador.
- Perry, J. 1975. *Personal Identity*: University of California Press.
- . 2003. *Knowledge, Possibility, and Consciousness*: MIT Press.
- Perry, John. 1972. "Can the Self Divide?" *The Journal of Philosophy* no. 69 (16):463-488.
- Piraino, Stefano, Ferdinando Boero, Brigitte Aeschbach, and Volker Schmid. 1996. "Reversing the Life Cycle: Medusae Transforming into Polyps and Cell

- Transdifferentiation in *Turritopsis nutricula* (Cnidaria, Hydrozoa)." *Biological Bulletin* no. 190 (3):302-312.
- Psillos, S. 2002. *Causation and explanation*: Acumen.
- Putnam, Hilary. 1973. "Meaning and Reference." *The Journal of Philosophy* no. 70 (19):699-711.
- Ramachandran, V.S. 2006. *A brief tour of human consciousness: from impostor poodles to purple numbers*: Pi Press.
- Rucker, Rudy v B. 2005. *Infinity and the mind : the science and philosophy of the infinite*. 2005 ed, *Princeton science library*. Princeton, N.J.: Princeton University Press.
- Sacks, O. 1986. *The man who mistook his wife for a hat*: Picador.
- Schechtman, M. 1996. *The constitution of selves*: Cornell University Press.
- Schroeter, Matthias L., Karolina Raczka, Jane Neumann, and D. Yves von Cramon. 2007. "Towards a nosology for frontotemporal lobar degenerations—A meta-analysis involving 267 subjects." *NeuroImage* no. 36 (3):497-510. doi: 10.1016/j.neuroimage.2007.03.024.
- Schwartz, Barry. 2009. *Our loss of wisdom*.
http://www.ted.com/talks/barry_schwartz_on_our_loss_of_wisdom.html.
- Shoemaker, S. 1984. "Personal Identity: A Materialist's Account." In *Personal identity*, edited by S. Shoemaker and R. Swinburne. B. Blackwell.
- Sider, T. 2001. *Four-dimensionalism: an ontology of persistence and time*: Clarendon Press.
- . 2010. *Is metaphysics about the real world? (Talk at California State University)*.
<http://www.youtube.com/watch?v=RKYZ8U-P5jA>.
- . 2012. *Writing the Book of the World*: Oxford University Press, Incorporated.
- Sider, T., J.P. Hawthorne, and D.W. Zimmerman. 2008. *Contemporary debates in metaphysics*: Blackwell Pub.
- Smart, J.J.C. 2008. "The Tenseless Theory of Time." In *Contemporary debates in metaphysics*, edited by T. Sider, J.P. Hawthorne and D.W. Zimmerman, 226-238. Blackwell Pub.
- Snowdon, Paul. 2003. "Objections to Animalism." In *On Human Persons (Metaphysical Research, Volume 1)*, edited by Klaus Petrus, 47-66. Ontos Verlag: Heusenstamm nr Frankfurt.
- Thomson, J.J. 2008. "People and Their Bodies." In *Contemporary debates in metaphysics*, edited by T. Sider, J.P. Hawthorne and D.W. Zimmerman, 155-176. Blackwell Pub.
- Wiggins, D. 1967. *Identity and spatio-temporal continuity*: Blackwell.
- Woit, P. 2006. *Not even wrong: the failure of string theory and the search for unity in physical law*: Basic Books.
- Zimmerman, D.W. 2008a. "The Privileged Present: Defending an "A-Theory" of Time." In *Contemporary debates in metaphysics*, edited by T. Sider, J.P. Hawthorne and D.W. Zimmerman, 211-225. Blackwell Pub.
- . 2008b. "Problems for animalism." *Abstracta* (Special Issue 1):23-31.

